

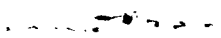
Fr m: Mertz, Prema
Sent: Thursday, January 24, 2002 10:24 AM
T : STIC-Biotech/ChemLib
Subject: 09/521,195

Please search SEQ ID NO:1 and 3 with protein databases.

Thanks
Prema Mertz, Ph.D.
Art Unit 1646
Mailbox 10C-01
Crystal Mall 1, Room 10E-01
United States Patent & Trademark Office
(703) 308-4229

Edward Hart
Technical Info Specialist
STIC / Biotech
CM1 12C14 Tel: 305-9203

TYPE OF SEARCH:		VENDOR/COST(where applic.)
Searcher: _____	NA Sequences: _____	STN: _____
Phone: _____	AA Sequences: <u>2</u>	DIALOG: _____
Location: _____	Structures: _____	Questel/Orbit: _____
Date Picked Up: <u>1/25/02</u>	Bibliographic: _____	DRLink: _____
Date Completed: <u>1/25/02</u>	Litigation: _____	Lexis/Nexis: _____
Searcher Prep/Review: _____	Full text: _____	Sequence Sys.: <u>02</u>
Clerical: _____	Patent Family: _____	WWW/Internet: _____
Online time: _____	Other: _____	Other (specify): _____



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on January 25, 2002, 10:08:19 ; Search time 43.11 Seconds
(without alignments)
946.750 Million cell updates/sec

US-09-521-195-1
score: 2845
e 1 MRDYDEVAFNGWGFQRL.....KKTRDSMTENPKVLITAF 551

, table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ied 522463 segs, 74073290 residues

number of hits satisfying chosen parameters: 522463

um DB seq length: 0

-processing: Minimum Match 0%

Listing first 45 summaries

abase : A_Geneseq_1101.*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than the actual score of the best hit. The results are sorted by descending predicted score and are derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	2845	100.0	551	20	AAV01649
2	2470	86.8	553	20	AAV01651
3	2470	86.8	553	21	AA820579
4	2218	78.0	557	20	AAV01650
5	2218	78.0	557	21	AA83929
6	2164	76.1	557	20	AA820580
7	2164	76.1	557	21	AA820580
8	2164	76.1	557	21	AA83930
9	2034	5	71	5	AA820578
10	779	5	27	4	AA843038
11	779	5	27	4	AA800930

12	698	24.5	540	22	AA849401
13	685.5	24.0	554	19	AAW64538
14	682	24.0	535	21	AAV51249
15	661.5	23.4	556	17	AA77676
16	664.5	23.3	546	22	AA847272
17	663.5	23.3	538	22	AA847273
18	652.5	22.0	545	21	AAV92902
19	652.5	22.0	545	21	AAV92902
20	650	22.8	551	21	AA808824
21	647	22.7	542	22	AA847274
22	637	22.4	536	21	AAV92903
23	636	22.4	561	18	AAW44196
24	630	22.1	537	18	AAW44195
25	624	21.8	537	18	AAW44195
26	621	21.8	578	22	AAE06571
27	614.5	21.5	483	22	AA800982
28	606.5	21.3	545	22	AA836553
29	597.5	21.0	551	20	AAW88488
30	586	20.6	550	21	AAV44278
31	583.5	20.5	520	22	AA847276
32	583.5	20.5	520	22	AA847276
33	564	19.8	550	22	AAE06612
34	564	19.8	550	22	AA869091
35	471.5	16.6	541	22	AA847276
36	430.5	15.1	553	22	AAE04897
37	417	14.7	480	21	AA828044
38	417	14.7	480	21	AA828044
39	412.5	14.5	521	21	AA847275
40	402	14.1	554	22	AA847275
41	387.5	13.6	397	21	AA830886
42	382.5	13.4	515	21	AA842522
43	341	12.0	360	21	AA830887
44	325.5	11.4	378	21	AA842523
45	312.5	11.0	359	21	AA842524

ALIGNMENTS

RESULT 1
AAV01649
ID AAV01649 standard; Protein: 551 AA.
AC AAV01649;
DT 23-JUN-1999 (first entry)
DE A protein with cation transporting activity.
KW Organic cation transporter; OCT1; OCT2; drug development; fatty liver, heart disease; cancer; anti-tumour drug; anticancer drug.
OS Homo sapiens.
XX WO9913072-A1.
XX PD 18-MAR-1999.
XX PF 07-SEP-1998; 98WO-JP04009.
XX PR 20-MAY-1999; 98JP-015650.
XX PR 08-SEP-1997; 97JP-0260972.
XX PA (CHUGAI) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX PI Nezu J, Oku A;
XX WP1. 1999-215063/18.
XX WP2. 1999-215063/18.
XX WP3. 1999-215063/18.
XX WP4. 1999-215063/18.
XX WP5. 1999-215063/18.
XX WP6. 1999-215063/18.
XX WP7. 1999-215063/18.
XX WP8. 1999-215063/18.
XX WP9. 1999-215063/18.
XX WP10. 1999-215063/18.
XX WP11. 1999-215063/18.
XX WP12. 1999-215063/18.
XX WP13. 1999-215063/18.
XX WP14. 1999-215063/18.
XX WP15. 1999-215063/18.
XX WP16. 1999-215063/18.
XX WP17. 1999-215063/18.
XX WP18. 1999-215063/18.
XX WP19. 1999-215063/18.
XX WP20. 1999-215063/18.
XX WP21. 1999-215063/18.
XX WP22. 1999-215063/18.
XX WP23. 1999-215063/18.
XX WP24. 1999-215063/18.
XX WP25. 1999-215063/18.
XX WP26. 1999-215063/18.
XX WP27. 1999-215063/18.
XX WP28. 1999-215063/18.
XX WP29. 1999-215063/18.
XX WP30. 1999-215063/18.
XX WP31. 1999-215063/18.
XX WP32. 1999-215063/18.
XX WP33. 1999-215063/18.
XX WP34. 1999-215063/18.
XX WP35. 1999-215063/18.
XX WP36. 1999-215063/18.
XX WP37. 1999-215063/18.
XX WP38. 1999-215063/18.
XX WP39. 1999-215063/18.
XX WP40. 1999-215063/18.
XX WP41. 1999-215063/18.
XX WP42. 1999-215063/18.
XX WP43. 1999-215063/18.
XX WP44. 1999-215063/18.
XX WP45. 1999-215063/18.
XX WP46. 1999-215063/18.
XX WP47. 1999-215063/18.
XX WP48. 1999-215063/18.
XX WP49. 1999-215063/18.
XX WP50. 1999-215063/18.
XX WP51. 1999-215063/18.
XX WP52. 1999-215063/18.
XX WP53. 1999-215063/18.
XX WP54. 1999-215063/18.
XX WP55. 1999-215063/18.
XX WP56. 1999-215063/18.
XX WP57. 1999-215063/18.
XX WP58. 1999-215063/18.
XX WP59. 1999-215063/18.
XX WP60. 1999-215063/18.
XX WP61. 1999-215063/18.
XX WP62. 1999-215063/18.
XX WP63. 1999-215063/18.
XX WP64. 1999-215063/18.
XX WP65. 1999-215063/18.
XX WP66. 1999-215063/18.
XX WP67. 1999-215063/18.
XX WP68. 1999-215063/18.
XX WP69. 1999-215063/18.
XX WP70. 1999-215063/18.
XX WP71. 1999-215063/18.
XX WP72. 1999-215063/18.
XX WP73. 1999-215063/18.
XX WP74. 1999-215063/18.
XX WP75. 1999-215063/18.
XX WP76. 1999-215063/18.
XX WP77. 1999-215063/18.
XX WP78. 1999-215063/18.
XX WP79. 1999-215063/18.
XX WP80. 1999-215063/18.
XX WP81. 1999-215063/18.
XX WP82. 1999-215063/18.
XX WP83. 1999-215063/18.
XX WP84. 1999-215063/18.
XX WP85. 1999-215063/18.
XX WP86. 1999-215063/18.
XX WP87. 1999-215063/18.
XX WP88. 1999-215063/18.
XX WP89. 1999-215063/18.
XX WP90. 1999-215063/18.
XX WP91. 1999-215063/18.
XX WP92. 1999-215063/18.
XX WP93. 1999-215063/18.
XX WP94. 1999-215063/18.
XX WP95. 1999-215063/18.
XX WP96. 1999-215063/18.
XX WP97. 1999-215063/18.
XX WP98. 1999-215063/18.
XX WP99. 1999-215063/18.
XX WP100. 1999-215063/18.
XX WP101. 1999-215063/18.
XX WP102. 1999-215063/18.
XX WP103. 1999-215063/18.
XX WP104. 1999-215063/18.
XX WP105. 1999-215063/18.
XX WP106. 1999-215063/18.
XX WP107. 1999-215063/18.
XX WP108. 1999-215063/18.
XX WP109. 1999-215063/18.
XX WP110. 1999-215063/18.
XX WP111. 1999-215063/18.
XX WP112. 1999-215063/18.
XX WP113. 1999-215063/18.
XX WP114. 1999-215063/18.
XX WP115. 1999-215063/18.
XX WP116. 1999-215063/18.
XX WP117. 1999-215063/18.
XX WP118. 1999-215063/18.
XX WP119. 1999-215063/18.
XX WP120. 1999-215063/18.
XX WP121. 1999-215063/18.
XX WP122. 1999-215063/18.
XX WP123. 1999-215063/18.
XX WP124. 1999-215063/18.
XX WP125. 1999-215063/18.
XX WP126. 1999-215063/18.
XX WP127. 1999-215063/18.
XX WP128. 1999-215063/18.
XX WP129. 1999-215063/18.
XX WP130. 1999-215063/18.
XX WP131. 1999-215063/18.
XX WP132. 1999-215063/18.
XX WP133. 1999-215063/18.
XX WP134. 1999-215063/18.
XX WP135. 1999-215063/18.
XX WP136. 1999-215063/18.
XX WP137. 1999-215063/18.
XX WP138. 1999-215063/18.
XX WP139. 1999-215063/18.
XX WP140. 1999-215063/18.
XX WP141. 1999-215063/18.
XX WP142. 1999-215063/18.
XX WP143. 1999-215063/18.
XX WP144. 1999-215063/18.
XX WP145. 1999-215063/18.
XX WP146. 1999-215063/18.
XX WP147. 1999-215063/18.
XX WP148. 1999-215063/18.
XX WP149. 1999-215063/18.
XX WP150. 1999-215063/18.
XX WP151. 1999-215063/18.
XX WP152. 1999-215063/18.
XX WP153. 1999-215063/18.
XX WP154. 1999-215063/18.
XX WP155. 1999-215063/18.
XX WP156. 1999-215063/18.
XX WP157. 1999-215063/18.
XX WP158. 1999-215063/18.
XX WP159. 1999-215063/18.
XX WP160. 1999-215063/18.
XX WP161. 1999-215063/18.
XX WP162. 1999-215063/18.
XX WP163. 1999-215063/18.
XX WP164. 1999-215063/18.
XX WP165. 1999-215063/18.
XX WP166. 1999-215063/18.
XX WP167. 1999-215063/18.
XX WP168. 1999-215063/18.
XX WP169. 1999-215063/18.
XX WP170. 1999-215063/18.
XX WP171. 1999-215063/18.
XX WP172. 1999-215063/18.
XX WP173. 1999-215063/18.
XX WP174. 1999-215063/18.
XX WP175. 1999-215063/18.
XX WP176. 1999-215063/18.
XX WP177. 1999-215063/18.
XX WP178. 1999-215063/18.
XX WP179. 1999-215063/18.
XX WP180. 1999-215063/18.
XX WP181. 1999-215063/18.
XX WP182. 1999-215063/18.
XX WP183. 1999-215063/18.
XX WP184. 1999-215063/18.
XX WP185. 1999-215063/18.
XX WP186. 1999-215063/18.
XX WP187. 1999-215063/18.
XX WP188. 1999-215063/18.
XX WP189. 1999-215063/18.
XX WP190. 1999-215063/18.
XX WP191. 1999-215063/18.
XX WP192. 1999-215063/18.
XX WP193. 1999-215063/18.
XX WP194. 1999-215063/18.
XX WP195. 1999-215063/18.
XX WP196. 1999-215063/18.
XX WP197. 1999-215063/18.
XX WP198. 1999-215063/18.
XX WP199. 1999-215063/18.
XX WP200. 1999-215063/18.
XX WP201. 1999-215063/18.
XX WP202. 1999-215063/18.
XX WP203. 1999-215063/18.
XX WP204. 1999-215063/18.
XX WP205. 1999-215063/18.
XX WP206. 1999-215063/18.
XX WP207. 1999-215063/18.
XX WP208. 1999-215063/18.
XX WP209. 1999-215063/18.
XX WP210. 1999-215063/18.
XX WP211. 1999-215063/18.
XX WP212. 1999-215063/18.
XX WP213. 1999-215063/18.
XX WP214. 1999-215063/18.
XX WP215. 1999-215063/18.
XX WP216. 1999-215063/18.
XX WP217. 1999-215063/18.
XX WP218. 1999-215063/18.
XX WP219. 1999-215063/18.
XX WP220. 1999-215063/18.
XX WP221. 1999-215063/18.
XX WP222. 1999-215063/18.
XX WP223. 1999-215063/18.
XX WP224. 1999-215063/18.
XX WP225. 1999-215063/18.
XX WP226. 1999-215063/18.
XX WP227. 1999-215063/18.
XX WP228. 1999-215063/18.
XX WP229. 1999-215063/18.
XX WP230. 1999-215063/18.
XX WP231. 1999-215063/18.
XX WP232. 1999-215063/18.
XX WP233. 1999-215063/18.
XX WP234. 1999-215063/18.
XX WP235. 1999-215063/18.
XX WP236. 1999-215063/18.
XX WP237. 1999-215063/18.
XX WP238. 1999-215063/18.
XX WP239. 1999-215063/18.
XX WP240. 1999-215063/18.
XX WP241. 1999-215063/18.
XX WP242. 1999-215063/18.
XX WP243. 1999-215063/18.
XX WP244. 1999-215063/18.
XX WP245. 1999-215063/18.
XX WP246. 1999-215063/18.
XX WP247. 1999-215063/18.
XX WP248. 1999-215063/18.
XX WP249. 1999-215063/18.
XX WP250. 1999-215063/18.
XX WP251. 1999-215063/18.
XX WP252. 1999-215063/18.
XX WP253. 1999-215063/18.
XX WP254. 1999-215063/18.
XX WP255. 1999-215063/18.
XX WP256. 1999-215063/18.
XX WP257. 1999-215063/18.
XX WP258. 1999-215063/18.
XX WP259. 1999-215063/18.
XX WP260. 1999-215063/18.
XX WP261. 1999-215063/18.
XX WP262. 1999-215063/18.
XX WP263. 1999-215063/18.
XX WP264. 1999-215063/18.
XX WP265. 1999-215063/18.
XX WP266. 1999-215063/18.
XX WP267. 1999-215063/18.
XX WP268. 1999-215063/18.
XX WP269. 1999-215063/18.
XX WP270. 1999-215063/18.
XX WP271. 1999-215063/18.
XX WP272. 1999-215063/18.
XX WP273. 1999-215063/18.
XX WP274. 1999-215063/18.
XX WP275. 1999-215063/18.
XX WP276. 1999-215063/18.
XX WP277. 1999-215063/18.
XX WP278. 1999-215063/18.
XX WP279. 1999-215063/18.
XX WP280. 1999-215063/18.
XX WP281. 1999-215063/18.
XX WP282. 1999-215063/18.
XX WP283. 1999-215063/18.
XX WP284. 1999-215063/18.
XX WP285. 1999-215063/18.
XX WP286. 1999-215063/18.
XX WP287. 1999-215063/18.
XX WP288. 1999-215063/18.
XX WP289. 1999-215063/18.
XX WP290. 1999-215063/18.
XX WP291. 1999-215063/18.
XX WP292. 1999-215063/18.
XX WP293. 1999-215063/18.
XX WP294. 1999-215063/18.
XX WP295. 1999-215063/18.
XX WP296. 1999-215063/18.
XX WP297. 1999-215063/18.
XX WP298. 1999-215063/18.
XX WP299. 1999-215063/18.
XX WP300. 1999-215063/18.
XX WP301. 1999-215063/18.
XX WP302. 1999-215063/18.
XX WP303. 1999-215063/18.
XX WP304. 1999-215063/18.
XX WP305. 1999-215063/18.
XX WP306. 1999-215063/18.
XX WP307. 1999-215063/18.
XX WP308. 1999-215063/18.
XX WP309. 1999-215063/18.
XX WP310. 1999-215063/18.
XX WP311. 1999-215063/18.
XX WP312. 1999-215063/18.
XX WP313. 1999-215063/18.
XX WP314. 1999-215063/18.
XX WP315. 1999-215063/18.
XX WP316. 1999-215063/18.
XX WP317. 1999-215063/18.
XX WP318. 1999-215063/18.
XX WP319. 1999-215063/18.
XX WP320. 1999-215063/18.
XX WP321. 1999-215063/18.
XX WP322. 1999-215063/18.
XX WP323. 1999-215063/18.
XX WP324. 1999-215063/18.
XX WP325. 1999-215063/18.
XX WP326. 1999-215063/18.
XX WP327. 1999-215063/18.
XX WP328. 1999-215063/18.
XX WP329. 1999-215063/18.
XX WP330. 1999-215063/18.
XX WP331. 1999-215063/18.
XX WP332. 1999-215063/18.
XX WP333. 1999-215063/18.
XX WP334. 1999-215

XX Claim 1, Page 41-45; 97pp: Japanese.
 XX The present sequence represents a protein with cation transporting
 CC activity. The genes are significantly homologous with organic cation
 CC transporters OCT1 and OCT2. The genes may be used in drug development,
 CC particularly in the treatment of diseases due to abnormality of the
 CC organic cation transporter functions e.g. fatty liver, heart diseases
 CC and cancers, by controlling such as by inhibition or activation.
 CC Administration of anti-tumour and anticancer drugs in combination with
 CC a transporter protein inhibiting agent allows the agents to penetrate
 CC into the diseased cells to enhance the drug action.
 XX Sequence 551 AA;
 XX

Query Match 100.0%; Score 2845; DB 20; Length 551;
 Best Local Similarity 100.0%; Pred. No. 1.5e-281; Indels 0; Gaps 0;
 Matches 551; Conservative 0; Mismatches 0;
 QY 1 MRDYDEVIAFLGEGWPGFQRIIFLLSASIPNGFNGMSVVFAGTPEHRCRVPDAANLSS 60
 Db 1 mrdydeviaflgwgpgfqrllfllsasilpngfngmsvvlagtpehrcrvpdaanlss 60
 QY 61 AWRNNSVPLRLDGRREVPHSCRYELATTANFSALEPGRDVGLQESCLDGEWFS 120
 Db 61 awrnnsvplrlldgrrevphscryelatnfsalglepgrdvldgleqescldgweifs 120
 QY 121 ODVYLSVVTENNLVCEENKVPVLTSLFFVGLVGSFVSGQLSDRFGRKNVLFATMAVQ 180
 Db 121 odvylsvvtennlvceenkvplvtslffvglvgsfsvsgqlsdrfgrknvlfatmavq 180
 QY 181 TGFSLQFQISWEMFVLVTVGNGOISNVVAFILGTILGKSVRIIPSTPLGVCTEFA 240
 Db 181 tgfslqfqiiswemfvlvtnvgngoisnvvafilgtilgksvriipstplgvcteffa 240
 Y 241 VGYMLLPAYFIKRWMLLLALTVPGLVCPVLPWFIPESPRWLSQRFEADIIQKA 300
 1 241 vgymlplpayfirkwmlalltvpglvcplvpwfipesprwlsqrfeadiiqlka 300
 301 AKNNNTAVPAVIFDSVE--ELNPLKOKAFILDLFRTNIAIMTSLMLMLTSVGYFALS 360
 301 aknnntavpavifdsveelpknplkokaflldlfrtniaimtsllmlmltsvgyfals 360
 361 LDAPNLHGDAYNCFPLSALIEIPAYITAWLLRLTPRRYIAAFLVGGGVLLFIQLVVP 420
 361 ldapnlhgddayncfplsalielpayitawllrltprryiaaflvgggvllfiqlvvp 420
 421 DYVSLTGLVLMGKGCITSAFSLVYTABLYPTLVNMAVGVTASTASVGSIIAPYFY 480
 421 dyvsltglvlmgkgcitfafslvytablyptlvnmavgvtastasvgsiiapyfy 480
 481 LGAYNMLPYVNGSLVLTIGFTLFPESLGMPLPETLEQMOKVWFRSGKTRDSMET 540
 481 lgaynmlpyvngslvltigftlfpeslgmplpetleqmokvwfrsgktrdsmet 540
 541 EENPKVLITAF 551
 41 eenpkvllitaf 551

2
 AAY01651 standard; Protein; 553 AA.
 AAY01651,
 23-JUN-1999 (first entry)
 A protein with cation transporting activity.
 Organic cation transporter; OCT1; OCT2; drug development; fatty liver;
 heart disease; cancer; anti-tumour drug; anticancer drug.
 . . .

XX Mus musculus.
 XX WO9913072-A1.
 XX 18-MAR-1999.
 XX 07-SEP-1998; 98WO-JP04009.
 XX 20-MAY-1998; 98JP-0156660.
 XX 08-SEP-1997; 97JP-0260972.
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX Nezu J, Oku A;
 XX WPI: 1999-215062/18.
 XX N-PSDB; AAX26898.
 PT Genes homologous with organic cation transporters OCT1 and OCT2,
 PT useful in design of new drugs for treatment of diseases due to
 PT abnormality of the transporter functions
 VS Claim 1; Page 63-67; 97pp: Japanese.
 CC The present sequence represents a protein with cation transporting
 CC activity. The genes are significantly homologous with organic cation
 CC transporters OCT1 and OCT2. The genes may be used in drug development,
 CC particularly in the treatment of diseases due to abnormality of the
 CC organic cation transporter functions e.g. fatty liver, heart diseases
 CC and cancers, by controlling such as by inhibition or activation.
 CC Administration of anti-tumour and anticancer drugs in combination with
 CC a transporter protein inhibiting agent allows the agents to penetrate
 CC into the diseased cells to enhance the drug action.
 XX Sequence 553 AA;
 XX

Query Match 86.8%; Score 2470; DB 20; Length 553;
 Best Local Similarity 84.6%; Pred. No. 2.9e-243;
 Matches 468; Conservative 39; Mismatches 44; Indels 2; Gaps 1;
 QY 1 MRDYDEVIAFLGEGWPGFQRIIFLLSASIPNGFNGMSVVFAGTPEHRCRVPDAANLSS 60
 Db 1 mrdydeviaflgwgpgfqrllfllsasilpngfngmsvvlagtpehrcrvpdtvnls 60
 QY 61 AWRNNSVPLRLDGRREVPHSCRYELATTANFSALEPGRDVGLQESCLDGEWFS 120
 Db 61 awrnnsvplrlldgrrevphscryelatnfsalglepgrdvldgleqescldgweifs 120
 QY 121 ODVYLSVVTENNLVCEENKVPVLTSLFFVGLVGSFVSGQLSDRFGRKNVLFATMAVQ 180
 Db 121 odvylsvvtennlvceenkvplvtslffvglvgsfsvsgqlsdrfgrknvlfatmavq 180
 Db 121 kdfistvtnwlvceadktpittslffvgvicsfvsqslsdrfgrkvvfatmavq 180
 QY 181 TGFSLQFQISWEMFVLVTVGNGOISNVVAFILGTILGKSVRIIPSTPLGVCTEFA 240
 Db 181 tgfslqfqiiswemfvlvtnvgngoisnvvafilgtilgksvriipstplgvcteffa 240
 Db 181 tgfslqfistnwmefvialvngmgqianvvafligtellskavriifstlgtvctffa 240
 QY 241 VGYMLLPAYFIKRWMLLLALTVPGLVCPVLPWFIPESPRWLSQRFEADIIQKA 300
 Db 241 vgymlplpayfirkwmlalltvpglvcplvpwfipesprwlsqrfeadiiqlka 300
 Db 241 igymvplpayfirkwmlalltvpglvcplvpwfipesprwlsqrfeadiiqlka 300
 QY 301 AKNNNTAVPAVIFDSVE--ELNPLKOKAFILDLFRTNIAIMTSLMLMLTSVGYFALS 358
 Db 301 aknnntavpavifdsveelpknplkokaflldlfrtniaimtsllmlmltsvgyfa 360
 QY 359 LSLDAPNLHGDAYNCFPLSALIEIPAYITAWLLRLTPRRYIAAFLVGGGVLLFIQLV 418
 Db 359 lslapnlhgddayncfplsalielpayitawllrltprryiaaflvgggvllfiqlv 420
 QY 419 PDVYFSLTGLVLMGKGCITSAFSLVYTABLYPTLVNMAVGVTASTASVGSIIAPYFY 478
 Db 419 pdvyfsltglvlmgkgcitfafslvytablyptlvnmavgvtastasvgsiiapyf 480
 Db 421 pedyfslvlgvmgkfgitfafslvytablyptlvnmavgvtastasvgsiiapyf 480

QY 479 VYLGAYNRMPLPYVMGSLVFLIGITFLPPESIGMTLPETLEOMQKVKWFRSGKKTDRSM 538
 Db 481 VYLGAYNRMPLPYVMGSLVFLIGITFLPPESIGMTLPETLEOMQKVKWFRSGKKTDRSM 540

QY 539 ETEENPKVLITAF 551
 Db 541 dreepkvlitaf 553

RESULT 3

AAB20579 AAB20579 standard; Protein; 553 AA.

XX AC AAB20579;

DT 11-DEC-2000 (first entry)

XX Mouse OCTN1 amino acid sequence.

DE Mouse; transporter; OCTN1; OCTN2; OCTN3; organic cation transporter;
 KW identification; regulator; carnitine transport.

XX OS Mus musculus.

PN W0200046368-A1.

XX 10-AUG-2000.

XX 04-FEB-2000; 2000WO-JP00619.

XX 05-FEB-1999; 99JP-0028406.

XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX Nezu J, Ose A;

XX WPI; 2000-586982/55.

XX Organic cation transporter gene OCTN3 expressed in testis for
 PT identification of regulators of carnitine transport for use as drugs

XX Example 2; Fig 2; 58pp; Japanese.

XX The present invention describes a mouse organic cation transporter
 CC protein (OCTN3). Also described are: (1) a method for screening
 CC compounds for their ability to regulate the transport of an organic
 CC cation into the cell, by generating a cell expressing OCTN3 at the
 CC cell membrane, contacting with the compound and organic cation, and
 CC observing the degree of transport of the organic cation; and (2) a
 CC method for screening compounds for their ability to be transported into
 CC the cell by OCTN3, by generating a cell expressing OCTN3 at the cell
 CC membrane, contacting with the compound and observing the degree of
 CC transport of the compound. OCTN3 can be used for the identification of
 CC regulators of the transport of organic cations (especially carnitine)
 CC into cells by OCTN3, for use as drugs. The present sequence represents
 CC the mouse OCTN1 amino acid sequence, which is used in an example from
 CC the present invention.

XX Sequence 553 AA;

Query Match 86.8%; Score 2470; DB 21; Length 553;
 Best Local Similarity 84.6%; Pred. No. 2.9e-243;
 Matches 468; Conservative 39; Mismatches 44; Indels 2; Gaps 1;

QY 1 MRDYDEVIAFLGEGPQRILFFLLSASIPNGFMGSMVVFVLAGTPEHRCRVPDANLSS 60
 Db 1 mrdydeviaflgpgpqrilffllsasiipngfmgsmsvfvlagtpehrcrlvpdtnlss 60

QY 61 AWRNNSVPLRLDGRVPHSCSYRLATTANFSALGLEPGRDVLGOLEOESCLDGEWS 120
 Db 61 swrnshp1e6kdgqrvgpccrryrlatiansamglepggdvleqieqscldgwey 120

QY 121 QDVYLSVTWTEWNLVCEDNKVKPLTJTSLEPVGVLGSGVSGQLSDRFGKKNVLPATMAVQ 180
 Db 121 kdlifstlvtewnlvceedwktpltsifvvgvicsfvgqlsdrfgrkknvfatmavq 180
 QY 181 TGFSFLQIFISISWEMFTVLFVIVMGQISNYYVAVFILGTEILGKSVRIIFSTLGVCTFFA 240
 Db 181 tgfsfvqlfstnwmftvlfvfaivgmqgisnyvavfildteilsksvriifstlvgctffa 240
 QY 241 VGYMLPLFAFYFIRDWRMLLALTVCVLCVPLWMTPESPRWLISORFREAEEDIQKA 300
 Db 241 igymvplfayfirdwrmllalticp9lfcvplwfpesprwlisqrfaeaeqika 300
 QY 301 AKMNTAVPAVIFDSVE--ELNPLKOOKAFILDLFTRNIAIMTINSLMLWLTSGVYFA 358
 Db 301 akmsivapagifdplelqelnsikqkvilddlfrtniatitvmavmlwmltsvgyfa 360
 QY 359 LSLDAPNLHGDAYLNCPLSALIEIPAYITAWLLRLTPRYITAAVLFWGGVLLFIOLV 418
 Db 361 lslnvpnlhgdvynclflsglievpayftawllrltpryiaagvltwggvllliqv 420
 QY 419 PVDYFSLGVLMLGKFGITSAFSLMYVFTAELYPTLVNRNAVGVITSTASRVGSIAPYF 478
 Db 421 pedyfsvsiglvmgkfgitsafsmlyvftaelypvlvrnmavgitsmasrvgsiapyf 480
 QY 479 VYLGAYNRMPLPYVMGSLVFLIGITFLPPESIGMTLPETLEOMQKVKWFRSGKKTDRSM 538
 Db 481 vylgaynrmplpyvmgslvfligittlffpesigvtilpenleqmkgvrgcrgkktvs 540
 QY 539 ETEENPKVLITAF 551
 Db 541 dreepkvlitaf 553

RESULT 4

AAY01650 AAY01650 standard; Protein; 557 AA.

XX AC AAY01650;

XX 23-JUN-1999 (first entry)

XX A protein with cation transporting activity.

XX Organic cation transporter; OCT1; OCT2; drug development; fatty liver;
 KW heart disease; cancer; anti-tumour drug; anticancer drug.

XX OS Homo sapiens.

XX W09913072-A1.

XX 18-MAR-1999.

XX 07-SEP-1998; 98WO-JP04009.

XX 20-MAY-1998; 98JP-0156660.

XX 08-SEP-1997; 97JP-0260972.

XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX Nezu J, Ose A;

XX WPI; 1999-215062/18.

XX N-PSDB; AAX26880.

PT Genes homologous with organic cation transporters OCT1 and OCT2,
 useful in design of new drugs for treatment of diseases due to
 PT abnormality of the transporter functions

PS Claim 1; Page 51-55; 97pp; Japanese.

XX The present sequence represents a protein with cation transporting

CC activity. The genes are significantly homologous with organic cation

CC transporters OCT1 and OCT2. The genes may used in drug development,
 CC particularly in the treatment of diseases due to abnormality of the
 CC organic cation transporter functions e.g. fatty liver, heart diseases
 CC and cancers, by controlling such as by inhibition or activation.
 CC Administration of anti-tumour and anticancer drugs in combination with
 CC a transporter protein inhibiting agent allows the agents to penetrate
 CC into the diseased cells to enhance the drug action.
 XX
 XX Sequence 557 AA;
 SQ

Query Match 78.0%; Score 2218; DB 20; Length 557;
 Best Local Similarity 75.9%; Pred. No. 1.5e-217;
 Matches 423; Conservative 57; Mismatches 71; Indels 6; Gaps 3;

QY 1 MRDYEVTAFLEGEWGPORLFFLLSASIPNGFNGMSVVFVLAGTPEHRCRVPDANLSS 60
 DB 1 mrdydevtaflgewpfgfllfllsasiipngftgssvflfatpehrcrvpdaanlss 60
 QY 61 AWRNNSVPLRLRDGEVPHSCSRYLATIANFSALGLEPGRVDLGLQEQESCLDGWEFS 120
 DB 61 awrnhtvplrlrdgrevphscrryrlatiansaalgldpgrdvdlggleqescldgwefs 120
 QY 121 QDVYLTSTVVTENLVNCEWKNVPLTTSIFVGVLLGSFVSQSLSDRFGKKNVLFATMAVQ 180
 DB 121 qdvyilstvvtvtenlvncedwnkvpלטלסלפגvlvgllgsfvsqslsdrfgkknvlfatvmq 180
 QY 181 TGFSELPQISISWENFTVLFVIVGMOISNVYVAFILGTILGKSVRIIFSTLGVTCTFFA 240
 DB 181 tgfslqifsknfemfvfvlvgmgqisnyvaafvltgilegksvriifstlgvcifya 240
 QY 241 VGYMLLPFAFIRDRWMLLALVPGVLCVPLWFWFIPESPRWLSIORFREAEIIOKA 300
 DB 241 fgymlplfayfirdwrmlvalmpgvlcvalwffipesprwlsiorfreaeiirka 300
 QY 301 AKMNTAVPAVIFD--SVEELNPLKQKAFILDLFRTRNTAINTMSLLMLLTSVGYFA 358
 DB 301 akangivvpstifdpseqlqdlsskkqgshnldlrltwnlrmvmtimsimlwmatisvgyfg 360
 QY 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRYIIAAVFWGGVGLFIQIV 418
 DB 361 lsldtgnlhgdifvncflsanvepayvlawlllqylprysmatalflggsvllfmqlv 420
 QY 419 PVDYFLSLGLVWLKFGITSAFSLMXYFTAEPLTVLRNMAVGVTSTASRGSIIAPYF 478
 DB 421 pbdylvlatvlvmvgkfgvtaafsmvvytaelyptvtrnmvgvgsstasrsglsilpyf 480
 QY 479 VYLGAYNRMPLVYVNGSLTVLIGITFTLFFPESLGMTLPETLEOMQKVKWFRSGK---KTR 535
 DB 481 vylgaydrflpylmgstiltaltiflpsfgtpltidqmlrvkgmkhrktpshtr 540
 QY 536 DSMETEENPKVL-ITAF 551
 DB 541 mldgqerptlikstaf 557

RESULT 5
 AAY83929
 ID AAY83929 standard; Protein; 557 AA.
 XX
 AC AAY83929;
 XX
 DT 05-JUL-2000 (first entry)
 XX
 DE Human carnitine transporter protein OCTN2.
 XX
 KW Organic cation transportation; human; carnitine transporter protein;
 KW OCTN2; diagnosis; systemic carnitine deficiency; mutation; gene therapy;
 KW Juvenile visceral steatosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200014210-A1.

XX 16-MAR-2000.
 PD
 XX 07-SEP-1999; 99WO-JP04853.
 PF
 XX 07-SEP-1998; 98JP-0252683.
 PR
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PA
 XX Nezu J, Oku A;
 PI
 XX WP1; 2000-256966/22.
 DR
 DR N-PSDB; AAA09889.
 XX
 XX Systemic carnitine deficiency gene OCTN2 encoding part of organic
 PT cation transporter, useful as diagnostic tool
 PT
 XX Example 1; Page 36-40; 106pp; Japanese.
 PS
 XX This sequence represents the human carnitine transporter protein OCTN2.
 CC The coding sequence can be used as a target for diagnosis of systemic
 CC carnitine deficiency by detecting the presence of mutations in the
 CC sequence, especially seen in the disease juvenile visceral steatosis
 CC (jvs). The wild type OCTN2 gene can be used in the gene therapy of the
 CC disease state.
 CC
 XX Sequence 557 AA;
 SQ

Query Match 78.0%; Score 2218; DB 21; Length 557;
 Best Local Similarity 75.9%; Pred. No. 1.5e-217;
 Matches 423; Conservative 57; Mismatches 71; Indels 6; Gaps 3;

QY 1 MRDYEVTAFLEGEWGPORLFFLLSASIPNGFNGMSVVFVLAGTPEHRCRVPDANLSS 60
 DB 1 mrdydevtaflgewpfgfllfllsasiipngftgssvflfatpehrcrvpdaanlss 60
 QY 61 AWRNNSVPLRLRDGEVPHSCSRYLATIANFSALGLEPGRVDLGLQEQESCLDGWEFS 120
 DB 61 awrnhtvplrlrdgrevphscrryrlatiansaalgldpgrdvdlggleqescldgwefs 120
 QY 121 QDVYLTSTVVTENLVNCEWKNVPLTTSIFVGVLLGSFVSQSLSDRFGKKNVLFATMAVQ 180
 DB 121 qdvyilstvvtvtenlvncedwnkvpלטלסלפגvlvgllgsfvsqslsdrfgkknvlfatvmq 180
 QY 181 TGFSELPQISISWENFTVLFVIVGMOISNVYVAFILGTILGKSVRIIFSTLGVTCTFFA 240
 DB 181 tgfslqifsknfemfvfvlvgmgqisnyvaafvltgilegksvriifstlgvcifya 240
 QY 241 VGYMLLPFAFIRDRWMLLALVPGVLCVPLWFWFIPESPRWLSIORFREAEIIOKA 300
 DB 241 fgymlplfayfirdwrmlvalmpgvlcvalwffipesprwlsiorfreaeiirka 300
 QY 301 AKMNTAVPAVIFD--SVEELNPLKQKAFILDLFRTRNTAINTMSLLMLLTSVGYFA 358
 DB 301 akangivvpstifdpseqlqdlsskkqgshnldlrltwnlrmvmtimsimlwmatisvgyfg 360
 QY 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRYIIAAVFWGGVGLFIQIV 418
 DB 361 lsldtgnlhgdifvncflsanvepayvlawlllqylprysmatalflggsvllfmqlv 420
 QY 419 PVDYFLSLGLVWLKFGITSAFSLMXYFTAEPLTVLRNMAVGVTSTASRGSIIAPYF 478
 DB 421 pbdylvlatvlvmvgkfgvtaafsmvvytaelyptvtrnmvgvgsstasrsglsilpyf 480
 QY 479 VYLGAYNRMPLVYVNGSLTVLIGITFTLFFPESLGMTLPETLEOMQKVKWFRSGK---KTR 535
 DB 481 vylgaydrflpylmgstiltaltiflpsfgtpltidqmlrvkgmkhrktpshtr 540
 QY 536 DSMETEENPKVL-ITAF 551
 DB 541 mldgqerptlikstaf 557

Db	241	fgfmvlpifayfirdwrmlaltvpqvlcgawlfipesprwlisggrikeaevirka	300
Qy	301	AKMNTAVPAVIED--SVEELNPLKQOKAFILDLFRTNIAIMTINSLILMLTSGYFA	358
Db	301	akingivapstifdpseidlnstkqhlhlydlrtmrivtimsiilwliisvyfg	360
Qy	359	LSDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRRYIIAALVFWGGVLLFLQLV	418
Db	361	lsldtpnlhgdlyvncflaavevpayvlawlllylpryysisaalfgsgvllfmqlv	420
Qy	419	PVDYFISIGLVMIGKFGITSFAFSLYVFAELYPVLVRNMAVGVTSASRVSILAPYF	478
Db	421	pselystalvmvgkfgitsaysmvvyvtaelyptvrnmvgvstasrlsilspfy	480
Qy	479	VYLGAYNRMLPYIVMGSITVLIGITLFFPESLGMTLPETLEOMQKVWFRSGK--KTR	535
Db	481	vylgaydrfpyilmsltlaltiffesfgvfpdpdtidqmlrvkgikwqigsqr	540
Qy	536	DSMETEENPKVL-ITAF 551	
Db	541	mqkgeesptvklstaf 557	
RESULT 7			
AA020580			
ID	AA020580	standard; Protein; 557 AA.	
XX	AA020580;		
AC	AA020580;		
XX	11-DEC-2000	(first entry)	
DT	11-DEC-2000	(first entry)	
DE	Mouse	OCTN2 amino acid sequence.	
XX	Mouse	OCTN2	
KW	Mouse; transporter; OCTN1; OCTN2; OCTN3; organic cation transporter;		
KW	identification; regulator; carnitine transport.		
XX	Mus musculus.		
OS	WO200046368-A1.		
PN	10-AUG-2000.		
XX	04-FEB-2000; 2000WO-JP00619.		
PF	05-FEB-1999; 99JP-0028406.		
PR	(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.		
XX	Nezu J, Ose A;		
PI	WPI; 2000-586982/55.		
XX	Organic cation transporter gene OCTN3 expressed in testis for		
DR	identification of regulators of carnitine transport for use as drugs		
XX	Example 2; Fig 2; 58pp; Japanese.		
PS	The present invention describes a mouse organic cation transporter		
CC	protein (OCTN3). Also described are: (1) a method for screening		
CC	compounds for their ability to regulate the transport of an organic		
CC	cation into the cell, by generating a cell expressing OCTN3 at the		
CC	cell membrane, contacting with the compound and organic cation, and		
CC	observing the degree of transport of the organic cation; and (2) a		
CC	method for screening compounds for their ability to be transported into		
CC	the cell by OCTN3, by generating a cell expressing OCTN3 at the cell		
CC	membrane, contacting with the compound and observing the degree of		
CC	transport of the compound. OCTN3 can be used for the identification of		
CC	regulators of the transport of organic cations (especially carnitine)		
CC	into cells by OCTN3, for use as drugs. The present sequence represents		
CC	the mouse OCTN3 amino acid sequence, which is used in an example from		
CC	the present invention.		
XX	Sequence	557 AA;	
SQ			

AA01652			
ID	AA01652	standard; Protein; 557 AA.	
XX	AA01652;		
AC	AA01652;		
XX	23-JUN-1999	(first entry)	
DT	23-JUN-1999	(first entry)	
DE	A protein with cation transporting activity.		
XX	Organic cation transporter; OCTN1; OCTN2; drug development; fatty liver;		
KW	heart disease; cancer; anti-tumour drug; anticancer drug.		
XX	Mus musculus.		
OS	WO9913072-A1.		
PN	18-MAR-1999.		
PD	07-SEP-1998; 98WO-JP04009.		
PF	20-MAY-1998; 98JP-0156660.		
PR	08-SEP-1997; 97JP-0260972.		
XX	(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.		
PA	Nezu J, Oku A;		
PI	WPI; 1999-215062/18.		
XX	N-PSDB; AAX26902.		
DR	Genes homologous with organic cation transporters OCT1 and OCT2,		
XX	useful in design of new drugs for treatment of diseases due to		
PT	abnormality of the transporter functions		
PT	Claim 1; Page 75-79; 97pp; Japanese.		
XX	The present sequence represents a protein with cation transporting		
CC	activity. The genes are significantly homologous with organic cation		
CC	transporters OCT1 and OCT2. The genes may be used in drug development,		
CC	particularly in the treatment of diseases due to abnormality of the		
CC	organic cation transporter functions e.g. fatty liver, heart diseases		
CC	and cancers, by controlling such as by inhibition or activation.		
CC	Administration of anti-tumour and anticancer drugs in combination with		
CC	a transporter protein inhibiting agent allows the agents to penetrate		
CC	into the diseased cells to enhance the drug action.		
XX	Sequence	557 AA;	
SQ			

Query Match	76.1%;	Score 2164;	DB 20;	Length 557;
Best Local Similarity	72.9%;	Pred. No. 4.9e-212;		
Matches	406;	Conservative	70;	Mismatches 75; Indels 6; Gaps 3;
Qy	1	MRDYDEVIAFLGEGPQRLFFLLSASIIIPNGFNMGSVVFLAGTPHRCRVDAANLSS	60	
Db	1	mrdydevtafigewgqrlifflisaiipngfngmsivflagtpchrcrvphtvnls	60	
Qy	61	AWRNNVPLRLDGRVPHSCSRYLATIANFSAIGLEPGRDVDLGOEESCLDGWEFS	120	
Db	61	awrnhsipletkdgrvpqkerryrlatiansfseiglepgrdvdlgeqescldgweyd	120	
Qy	121	ODVYLSTVVTWNLCEDNWKVPLTTSFFVGLLGSFVSGQLSDRFGRKNVLFATMAVQ	180	
Db	121	kdvflstivtewdlvckdwkapitstlffvgvlmgfsiqslsdrgrknvfltagmq	180	
Qy	181	TGFSFLQIFSISWEMFTLVFVVGQISYVAVFILGTEILGKSVRIIPSTLGVCTFFFA	240	
Db	181	tgfsflgvfsvnfemftlvflvlgmgqisnyvaafvlgteillsksirifatlvgcfya	240	
Qy	241	VOYMLLPFAFIRDRWMLLALVPGVLCVPLWFWIPESPRWLISQRFREAEIIQKA	300	

```

Query Match          76.1%; Score 2164; DB 21; Length 557;
Best Local Similarity 72.9%; Pred. No. 4.9e-212;
Matches 406; Conservative 70; Mismatches 75; Indels 6; Gaps 3;

Qy 1 MRDYDEVIAFLGEGWGPORLIFFLLSASIIIPNGFNGSVVFLAGTPEHRCRVPDAANLSS 60
Db 1 mrdydeviaflgwgwpqrllffllsasiipngfngmsvflagtpchrcrlvphntvnlss 60

Qy 61 AWRNNSVPLRLDRGVRPHSCSRRLATIANFSALEPGRDVDLQLEQESCLDGEWFS 120
Db 61 awrnhspletkdgrvqpkrryrlatianselglepdrvdleqlesclgdwey 120

Qy 121 QDVYVLTSTVTEWNLVCEENKVKPLTSLFFVGLGSFVSQGLSDRFRGNVLFATMAVQ 180
Db 121 kvdfistivtewdlvckdwdkapltslffvglmgsfvsqglssdrfrgnvlfatmgmq 180

Qy 181 TGFSLQIFSIWEMFTVLFVIVGMQISNVYVAFILGTTELKSKSVRIIFSTLGVCTFFA 240
Db 181 tgfslqifsvnfemftvlfvlgmgqisnyvaafvlgteillsksiriffatlgvcifva 240

Qy 241 VGYMLLPFAFIRDRMMLLALVPGVLCVPLWFWFIPESPRWLSIORRPREAEDIIOKA 300
Db 241 fgmvlplfafirdrwmlallatvpgvlgcplwfwfipesprwlsgrgriksaeavirka 300

Qy 301 AKMNTAVPAVIFD--SVEELNPLKQOKAFILDLFRTNIAIMTMSLLMLTSLVGYFA 358
Db 301 akingivapstifdpseqlqinstkplhhiydlirtnirvritmsiilwlcisvgyf 360

Qy 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRYITAAVLFWGGVLLFTQLV 418
Db 361 lslatpnlhgdvynclflaavevpayvawlllqypryysaalfggsvllfmqlv 420

Qy 419 PVDYVFLSIGLVMLGKFGITSAFSLMYVFAEYPTLVNRNAVGVSTASRVGSIIAPYF 478
Db 421 pselftyistalvmvgkfgitsaysmvvytaelyptlvnrnmvgvstasrllsipyf 480

Qy 479 VYLGAYNRLPYVNGSLTILGIFTLFFPESLGMTLPETLEOMOKVKWFRSK---KTR 535
Db 481 vylgaydrflpylmgsltiltaltlffpesfgvplpdtidqmlrvkgikwqidsqtr 540

Qy 536 DSMETEENPKVL-ITAF 551
Db 541 mqkgeesptvltkstaf 557

RESULT 8
AAY8930
ID AAY8930 standard; Protein: 557 AA.
XX
AC AAY8930;
XX
DT 05-JUL-2000 (first entry)
XX
DE Mouse carnitine transporter protein OCTN2.
XX
KW Organic cation transportation; mouse; carnitine transporter protein;
KW OCTN2; diagnosis; systemic carnitine deficiency; mutation; gene therapy;
KW juvenile visceral steatosis.
OS Mus musculus.
XX
PN WO200014210-A1.
XX
PD 16-MAR-2000.
XX
PF 07-SEP-1999; 99WO-UP04853.
XX
PR 07-SEP-1998; 98JP-0252683.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX

```

```

PI Nezu J, Oku A;
XX
DR WPI: 2000-256966/22.
DR N-PSDB: AAA09890.
XX
PT Systemic carnitine deficiency gene OCTN2 encoding part of organic
PT cation transporter, useful as diagnostic tool.
PS
XX Example 1; Page 47-51; 106pp; Japanese.
XX
CC This sequence represents the mouse carnitine transporter protein OCTN2.
CC The coding sequence of the corresponding human protein can be used as a
CC target for diagnosis of systemic carnitine deficiency by detecting the
CC presence of mutations in the sequence, especially seen in the disease
CC juvenile visceral steatosis (jvs). The wild type OCTN2 gene can be used
CC in the gene therapy of the disease state.
XX
SQ Sequence 557 AA;

Query Match          76.1%; Score 2164; DB 21; Length 557;
Best Local Similarity 72.9%; Pred. No. 4.9e-212;
Matches 406; Conservative 70; Mismatches 75; Indels 6; Gaps 3;

Qy 1 MRDYDEVIAFLGEGWGPORLIFFLLSASIIIPNGFNGSVVFLAGTPEHRCRVPDAANLSS 60
Db 1 mrdydeviaflgwgwpqrllffllsasiipngfngmsvflagtpchrcrlvphntvnlss 60

Qy 61 AWRNNSVPLRLDRGVRPHSCSRRLATIANFSALEPGRDVDLQLEQESCLDGEWFS 120
Db 61 awrnhspletkdgrvqpkrryrlatianselglepdrvdleqlesclgdwey 120

Qy 121 QDVYVLTSTVTEWNLVCEENKVKPLTSLFFVGLGSFVSQGLSDRFRGNVLFATMAVQ 180
Db 121 kvdfistivtewdlvckdwdkapltslffvglmgsfvsqglssdrfrgnvlfatmgmq 180

Qy 181 TGFSLQIFSIWEMFTVLFVIVGMQISNVYVAFILGTTELKSKSVRIIFSTLGVCTFFA 240
Db 181 tgfslqifsvnfemftvlfvlgmgqisnyvaafvlgteillsksiriffatlgvcifva 240

Qy 241 VGYMLLPFAFIRDRMMLLALVPGVLCVPLWFWFIPESPRWLSIORRPREAEDIIOKA 300
Db 241 fgmvlplfafirdrwmlallatvpgvlgcplwfwfipesprwlsgrgriksaeavirka 300

Qy 301 AKMNTAVPAVIFD--SVEELNPLKQOKAFILDLFRTNIAIMTMSLLMLTSLVGYFA 358
Db 301 akingivapstifdpseqlqinstkplhhiydlirtnirvritmsiilwlcisvgyf 360

Qy 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRYITAAVLFWGGVLLFTQLV 418
Db 361 lslatpnlhgdvynclflaavevpayvawlllqypryysaalfggsvllfmqlv 420

Qy 419 PVDYVFLSIGLVMLGKFGITSAFSLMYVFAEYPTLVNRNAVGVSTASRVGSIIAPYF 478
Db 421 pselftyistalvmvgkfgitsaysmvvytaelyptlvnrnmvgvstasrllsipyf 480

Qy 479 VYLGAYNRLPYVNGSLTILGIFTLFFPESLGMTLPETLEOMOKVKWFRSK---KTR 535
Db 481 vylgaydrflpylmgsltiltaltlffpesfgvplpdtidqmlrvkgikwqidsqtr 540

Qy 536 DSMETEENPKVL-ITAF 551
Db 541 mqkgeesptvltkstaf 557

RESULT 9
AAB20578
ID AAB20578 standard; Protein: 564 AA.
XX
AC AAB20578;
XX
DT 11-DEC-2000 (first entry)
XX

```

DE Mouse OCTN3 protein SEQ ID NO:1.
 XX Mouse; transporter; OCTN1; OCTN2; OCTN3; organic cation transporter;
 KW Identification; regulator; carnitine transport.
 XX Mus musculus.
 OS
 XX WC2000046368-A1.
 PN 10-AUG-2000.
 XX
 XX 04-FEB-2000; 2000WO-JP00619.
 XX 05-FEB-1999; 99JP-0028406.
 PR
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 PA
 XX Nezu J, Ose A;
 PI
 XX WPI: 2000-586982/55.
 DR N-PSDB; AAA88053.
 XX
 XX Organic cation transporter gene OCTN3 expressed in testis for
 PT identification of regulators of carnitine transport for use as drugs
 XX
 XX Claim 1; Page 34-39; 58pp; Japanese.
 XX
 CC The present invention describes a mouse organic cation transporter
 CC protein (OCTN3). Also described are: (1) a method for screening
 CC compounds for their ability to regulate the transport of an organic
 CC cation into the cell, by generating a cell expressing OCTN3 at the
 CC cell membrane, contacting with the compound and organic cation, and
 CC observing the degree of transport of the organic cation; and (2) a
 CC method for screening compounds for their ability to be transported into
 CC the cell by OCTN3, by generating a cell expressing OCTN3 at the cell
 CC membrane, contacting with the compound and observing the degree of
 CC transport of the compound. OCTN3 can be used for the identification of
 CC regulators of the transport of organic cations (especially carnitine)
 CC into cells by OCTN3, for use as drugs. The present sequence represents
 CC mouse OCTN3.
 XX
 SQ Sequence 564 AA;
 Query Match 71.5%; Score 2034.5; DB 21; Length 564;
 Best Local Similarity 69.08; Pred. No. 8.3e-199;
 Matches 380; Conservative 75; Mismatches 85; Indels 11; Gaps 3;
 QY 1 MRDVEVIAFLGEGPQFORLIFLLSASIIIPNGMSVWVFLAGTPEHRCRVPDAAALSS 60
 DB 1 mldydevtaflgwgftgrllfillsasilpngftgisavfitaiphehrcrptdvnlss 60
 QY 61 AWRNNSVPLRLRDGEVPHSCSRVRLATIANFSALGLEPGRDVDLGOLESCLDGWEFS 120
 DB 61 awrnhsipmetkddpvpqkrryrlatianfse.lgpegrdvldleqencldgweyd 120
 QY 121 ODVYLSVTWVTEWNLVCENWVPLTTSFFVGLLGSFVSGQLSDRFGRKNVLFATMAVQ 180
 DB 121 kdlifstivtedwlvckdwkaplttsffvvgvllgsfsgqlsdrgrknllftmamh 180
 QY 181 TGFSLQIFISWEMFTLVFVVGQISNYVAVFILGTEILGKSVRIIFSTLGVCFFFA 240
 DB 181 tgfslqifiswemftlvtlvvgmghisnyvaafvlgtemlsksvrliatfalgvciffa 240
 QY 241 VCYMLLPFAFIRDRWMLLALVPGVLCVPLWFWFPESPRWLLSQRPRFEADITQKA 300
 DB 241 tgfmlpfpafirwrrllailtpgvlcgallwfwfpeprwllsqgrikeaeiirka 300
 QY 301 AKMNNATPAVFDSDVEELNPL-----KQKAFILDLFRNTAINTMSLLMLTSTV 354
 DB 301 akngivapstfop-setnklqddskkqshhdydvtrcpnrlitmsiilwltisv 359
 QY 355 GYFALSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRRIIAAFLWGGVLLF 414

DB 360 gyfgisldtpnlnngniyvcncllaavepayvclawlllqhvsrrysmagsiflgsvlll 419
 QY 415 IOLVPDYDFYLSIGLVMLGKFGITSAFSLMYFTAEIYPTLVRNMAVGVTSASRVGSII 474
 DB 420 vqlvpsdlhyisttlvmvgkfgitsaysmvvytaelyptvrmvgvgsstasrlgsil 479
 QY 475 APFYVYLGAYNRMPLPIVVMGSLTVLIGITFTFFPESLGMTLPETLEQMKQVWFRSGRKT 534
 DB 480 spfyfylvaydrllpyllmgsiltaitiffpessgysipetidmqvkk1----kq 535
 QY 535 RDSMETEENPK 545
 DB 536 rqlskskkgspek 546
 RESULT 10
 AAB43038
 ID AAB43038 standard; Protein; 560 AA.
 XX
 AC AAB43038;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF2802 polypeptide sequence SEQ ID NO:5604.
 XX
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipariatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive.
 XX
 OS Homo sapiens.
 XX
 XX WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 XX 31-MAR-2000; 2000WO-US08621.
 XX
 XX 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkets RA, Leach M;
 XX WPI: 2000-602362/57.
 DR N-PSDB; AAC77247.
 DR
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 4791-4792; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
 CC antipariatic; antiparkinsonian; nootropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;

CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 XX Sequence 560 AA;
 Query Match 27.4%; Score 779.5; DB 21; Length 560;
 Best Local Similarity 32.9%; Pred. No. 9.8e-71;
 Matches 184; Conservative 113; Mismatches 214; Indels 49; Gaps 9;
 QY 17 FORLIFLLSASTIPNGFMGMSVFLAGTPEHRCRVP-----DAANLS 59
 Db 2 fqrlyfcafnisqihylasvfmgtphhvcprpqnsvqvfhnhsnwledtgall 61
 QY 60 SAWNNNSVPLRLDRCREVPHS-CSRYRLATIANFSAIGLEPRDVLGQLESCLDGWE 118
 Db 62 ssgqdyvtvqlngelwelsrcnkr-----entslgye-----ytskkefcpcvdyi 113
 QY 119 FSQDYLSTVVTWEMNLVCEDDNKVPLTTLSLFFVGLLGSFVSQGLSDRGRKNVLFATMA 178
 Db 114 ydntkwtstavgtnlvcdrkwlamlqplfmfgvllgstvgfysdrlgrvrvlwatss 173
 QY 179 VQTGFSFLQIFSIWEMFVFLFVIVGMGOISNVYVAFILGTEILGKSVRIIFSTLGVCTF 238
 Db 174 smflglaafavdytffmaarfflamvasgylvvgvymfegmkst-twasvhlhsf 232
 QY 239 FAVGYMLLPFAFYFIDW---RMLLALTVPGVLCVPLWVFPEPRWLSORPREAED 295
 Db 233 favgtlilvaltgylyrtwlygmilstvvtvfilcc---wlpfcplwllsegryeesqk 289
 QY 296 ITOKAAMKNTA-----VPAVIFDSVEELNPKOOKAFILDLFRTRNIAIMTMSLLW 349
 Db 290 ivdimakwnrasscklsellslldgppvsnspetevqkhnslvfyfnwsitkrtltvwlw 349
 QY 350 MLTSVGYFALSIDAPNLHGDAYLNGFLSALIEIPAYITAWLLRLTPRRIIAAVLFWGG 409
 Db 350 ftgsigfysfnsnlgneylnflgvgveipaytfvclamdkgvrtvlaylsfcsa 409
 QY 410 GVLLFTQLVPDVPYFLISIGLVMLGFGITSAFMSLYVFTAEYLPVLVRNMAVGVSTASR 469
 Db 410 lacgvvmvipqkhyllgvvtamvgfagaafglylytaelyptiyrslavsgsmvcr 469
 QY 470 VGSITAPYFVYLCANRMPLPYVIMGSLVNLVIGTIFLFFPESLGWTLPTLEOMOKVKNFR 529
 Db 470 lasilaptsvdlsswifipqlfvgtmallsgvltlkipetlgrlattweeaakl---- 525
 QY 530 SGKTRDSMETEENPKVLT 549
 Db 526 -----esenesksklllt 539
 RESULT 11
 AAM00930
 ID AAM00930 standard; Protein; 584 AA.
 XX
 AC AAM00930;
 XX
 DT 01-OCT-2001 (first entry)
 XX
 DE Human bone marrow protein, SEQ ID NO: 406.

XX
 KW Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KW immunosuppressive; gene therapy; cytokine cell proliferation;
 KW cell differentiation modulator; immune disorder; infection; cancer;
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
 XX
 OS Homo sapiens.
 XX WO200153453-A2.
 XX 26-JUL-2001.
 XX 23-DEC-2000; 2000WO-US34960.
 XX 21-JAN-2000; 2000US-0488725.
 XX 25-APR-2000; 2000US-0552317.
 XX 09-JUL-2000; 2000US-0598042.
 XX 19-JUL-2000; 2000US-0620312.
 XX 03-AUG-2000; 2000US-0653450.
 XX 14-SEP-2000; 2000US-0662191.
 XX 19-OCT-2000; 2000US-0693036.
 XX 30-NOV-2000; 2000US-0250583.
 XX (HYSE-) HYSEQ INC.
 XX Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 XX Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 XX Zhao QA, Zhou P, Drmanac RT;
 XX WPI: 2001-488707/53.
 XX N-PSDB: AAH90049.
 XX Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 XX for treating e.g. cancer and immune deficiency disorders -
 XX Claim 10; Page 490-491; 648pp; English.
 CC The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX Sequence 584 AA;
 Query Match 27.4%; Score 779; DB 22; Length 584;
 Best Local Similarity 32.3%; Pred. No. 1.2e-70;
 Matches 186; Conservative 118; Mismatches 222; Indels 50; Gaps 10;
 QY 2 RYDEVITAFGEWGPQRLIFLLSASIIIPNGFMGMSVFLAGTPEHRCRVP----- 53
 Db 10 rhfegiydhvghrgrfgrvlyfcafnisqihylasvfmgtphhvcprpqnsvqvf 69
 QY 54 -----DAANLSAARNNSVPLRLDRCREVPHS-CSRYRLATIANFSAIGLEPRDV 103
 Db 70 hnhsnwledtgallssgqdyvtvqlngelwelsrcnkr-----entslgye----- 121
 QY 104 DLGQLEQESCLDGWEFSQDYLSTVVTWEMNLVCEDDNKVPLTTLSLFFVGLLGSFVSQGL 163
 Db 122 ytskkefcpcvdyi ydntkwtstavgtnlvcdrkwlamlqplfmfgvllgstvgf 181
 QY 164 SDRFGRKNVLFATMAVOTGFSFLQIFSIWEMFVFLFVIVGMGOISNVYVAFILGTEILG 223

```

182 sdrlgrrvltwssmflgfaafaydytfaarfllamvasgylvvgfvyvmefig 241
224 KSVRIIFSLGCTFFAVGYMLLPFLAYFIRDW-----RMLLALTPVGLVPLMWFIPES 280
242 mksr-twasvhlhsffavgtllvaltgylvrtwlygmilstvtvpfilcc---wvipet 297
281 PRLWLSQRPRAEADIIQKAAMNNTA-----VPAVIFDSVEELNPLKQKAFILDLFR 334
298 pfwllsegryeeaqkivdimakwnrascklsellisdlqgvnsnpcevkahnlsylfy 357
335 TRNIAMTMSLLMLTSVGVFALSLDAPNLHGDAYLNCFLSALIEIPAYITAWMLLRT 394
358 nwsikrtltvwlftsglsgfslsvnlgnneylnflilgveipaytfvciamdk 417
395 LPRYITIAAVLFWGGVLLFIQLVPDYFYLISGLVGLK-FGITSAPSMLYVFTAEIYP 453
418 vgrtviaylsfcsalacgvnmvipqkhyilgvvtamvgkilpigaafglylytaelyp 477
454 TLVRNMAVGVTASRGVSTIAPFYVILGAYNKMPLPYVMGSLTVLIGITLFFPESLGM 513
478 tlrvslavsgsmvcrilasilapfsvdissiwifipqlfvgtmallsgvltklpetlgk 537
514 TLPETLEQMQKWFRRSGKTKRDSMETENPKVLIT 549
538 rlatwteeaakl-----eseneskskillit 563

RESULT 12
AAB49401
ID AAB49401 standard; Protein; 540 AA.
AC AAB49401;
XX
XX
DT 02-MAR-2001 (first entry)
DE Murine organic anion transporter 6.
KW Murine; organic anion transporter 6; mOATP6; cancer; inflammation;
KW cardiovascular disease; central nervous system disorder; kidney disease;
KW liver disease; autoimmune disease.
XX Mus sp.
XX W0200070048-A1.
XX 23-NOV-2000.
XX 15-MAY-2000; 2000WO-US13316.
XX 14-MAY-1999; 99US-0134137.
XX 12-MAY-2000; 2000US-0570293.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX Feild J, Yue L, Ellens H;
XX WPI; 2001-016235/02.
XX N-PSDB; AAC83979.
XX Murine organic anion transporter 6 polypeptide, useful for identifying
XX agonists/antagonists that are useful in treatment of cancer, kidney
XX disease, autoimmune disease, inflammation and cardiovascular disease
XX Claim 2; Page 28-29; 32pp; English.
XX The present sequence is murine organic anion transporter 6 (mOATP6).
XX mOATP6 protein is useful for screening compounds which inhibit or
XX stimulate the function of mOATP6 and also compounds that neither agonise
XX nor antagonise OATP6. The identified agonists and antagonists are useful
XX for prevention and treatment of human diseases, including cancer,
XX inflammation, cardiovascular disease, central nervous system disorders,

```

CC kidney diseases, liver disease and autoimmune diseases.

XX Sequence 540 AA;

Query Match 24.5%; Score 698; DB 22; Length 540;
 Best Local Similarity 32.5%; Pred. No. 1.9e-62;
 Matches 183; Conservative 97; Mismatches 231; Indels 52; Gaps 19;

```

QY 4 YDEVIAFLGEGWGFQ--RLIFFLLSASIIIPNGNGSVVFLAGTPEHRCRVPDA-ANLS- 59
DB 3 feelihkvvgfpgfqlnvlalprflipmhf--lipfmaavpahhcaipapanlsh 60
QY 60 -SARNNSVPLRLRDGREVPVPHSCSRYLATIANFSALGLEPGRDVLGLQEQE-----SCL 114
DB 61 qdlwlkthlp-retdgsf--ssclrfaypqalpnvtlgtv---ynsgepegepltpcs 114
QY 115 DGWEFSODVYLSTVVTENLVCEENWKVPITSLTFFVGVLLGSFVSGQLSDRGKKNWLF 174
DB 115 qgweydrsefsstiatcdwlcqgrglnkvtstcffiigllgavvyeylsdrfrrill 174
QY 175 ATMAVQTGFSFLQIFISWEMFTVLFVIVGMQISNVVAFILGTEILGKSVRIIFSTLG 234
DB 175 vayvstlalglmsasvnylmfvtrmltgsalagftliivlplewldvehrtvagvis 234
QY 235 VCTFFAVGYMLLPFLAYFIRDWRMLLALTPVGLVCPV---LWMFIPESPRMLISORRF 290
DB 235 t-tftvgvllltlvgylirswrwillaatlp---cypgiisiww--vpesarwlltgrv 289
QY 291 REAEDIITQAKAMNNTAVPAVIFDSV--EELNPL-----KQKAFILDLFETRNIAIMI 343
DB 290 eeakkyisicaklngripse---dlsqsealnkvitmervsqrpsylidfrtsqlrhvsl 346
QY 344 MSLLLMULTSVGYFALSLDAPNLHGDAYLNCFLSALIEIPAYITAWMLLRLTPRR----- 398
DB 347 commmwfgnfsyvgltldasgltyqtllfgavevpskitvfflrvlgrlreag 406
QY 399 YIIAAVLFWGGVLLFIQLVPDYFYLISGLVGLKFGITSAPSMLYVFTAEIYPTLVN 458
DB 407 mllatalctfgisl-----lvssdtkswitalvwigkafseaaftaylftselypvlrg 461
QY 459 MAVGVTSASRGVSTIAPFYVILGAYNKMPLPYVMGSLTVLIGITLFFPESLGMTLPET 518
DB 462 tgmgtfalgirgaslapvllldgwwlllpklayggisflaactvlllpetkkaipet 521
QY 519 LEQMOKVKWFRSGKTKRDSMETE 541
DB 522 lqdv-----rkgrki-drsqte 538

RESULT 13
AAB4538
ID AAB4538 standard; Protein; 554 AA.
XX
XX AAB4538;
XX
XX 21-OCT-1998 (first entry)
XX Human liver cell clone HP01293 protein.
XX Transmembrane domain; human; nutrition; cytokine; cell proliferation;
XX differentiation; immune system; stimulator; suppressor; regulator;
XX hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;
XX haenostatic; thrombolytic; ligand; anti-inflammatory; tumour.
XX Homo sapiens.
XX WO9821328-A2.
XX 22-MAY-1998.
XX 07-NOV-1997; 97WO-JP04056.
XX

```



```

Db 408 ll--gaaltfgtllvsletkswitalvvvgkafseaafttaylftselyptvltqtlg 465
Qy 463 VTASRVSIIAPYVYIGAYNRMLPYIVMGSLTVLIGIFILFPSPSLGWLTPETLEOM 522
Db 466 ltalmgrlgaslaplaalldgwlilpkvayggaiaaactallipetkkaqlpetigdv 525
Qy 523 QK 524
Db 526 er 527

RESULT 15
AAR77676
ID AAR77676 standard; Protein; 556 AA.
XX
AC AAR77676;
XX
DT 10-SEP-1996 (first entry)
XX
DE Rat OCT-1 protein.
XX
KW Rat; OCT-1; transporter protein; cationic; xenobiotic; pharmaceutical;
KW blood; liver; kidney; epithelial cell; intestine; tetraethylammonium;
KW proximal renal tubule cell; intestine; enterocyte; transgenic; renal;
KW biliary; excretion; resorption; modulator; uptake.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT Domain 20..46
FT /note= "transmembrane domain"
FT Modified-site 71
FT /note= "putative N-glycosylation site"
FT Modified-site 97
FT /note= "putative N-glycosylation site"
FT Modified-site 113
FT /note= "putative N-glycosylation site"
FT Domain 154..171
FT /note= "transmembrane domain"
FT Domain 178..197
FT /note= "transmembrane domain"
FT Domain 243..260
FT /note= "transmembrane domain"
FT Domain 267..283
FT /note= "transmembrane domain"
FT Domain 350..366
FT /note= "transmembrane domain"
FT Domain 380..398
FT /note= "transmembrane domain"
FT Domain 406..425
FT /note= "transmembrane domain"
FT Modified-site 432
FT /note= "putative N-glycosylation site"
FT Domain 435..452
FT /note= "transmembrane domain"
FT Domain 469..485
FT /note= "transmembrane domain"
FT Domain 494..514
FT /note= "transmembrane domain"
XX
PN DE4424577-A1.
XX
PD 18-JAN-1996.
XX
PF 13-JUL-1994; 94DE-4424577.
XX
PR 13-JUL-1994; 94DE-4424577.
XX
PA (FARH ) HOECHST AG.
XX
PI Gorboulev V, Gruendeman D, Koepsell H;
XX

```

```

DR WPI; 1996-069422/08.
DR N-PSDB; AAT08702.
XX
PT Transporter protein for cationic xenobiotic(s) and pharmaceuticals,
PT and related DNA and transformed cells - used e.g. to assess
PT excretion and resorption of cationic cpds.
XX
PS Claim 1; Fig 2A1; 13pp; German.
XX
CC This is the amino acid sequence of the rat OCT-1 protein, a new
CC transporter protein able to transport cationic xenobiotics and
CC pharmaceuticals from the blood into liver or kidney epithelial
CC cells or from the intestine. The gene was isolated by injecting
CC a rat kidney gene library into Xenopus laevis oocyte and isolating
CC clones conferring uptake of 14C-tetraethylammonium. One clone
CC designated OCT-1 was isolated. Expression of the gene was detect
CC in proximal renal tubule cells, in liver epithelial cell and in
CC intestinal enterocytes. The DNA can be used to generate transgenic
CC cells for use in vitro test for renal/biliary excretion or
CC intestinal resorption of xenobiotics and pharmaceuticals. The protein
CC or cells expressing it can also be used to isolate modulators that
CC block uptake of pharmaceutical by the renal tubules.
XX
SQ Sequence 556 AA;
Query Match 24.0%; Score 681.5; DB 17; Length 556;
Best Local Similarity 33.0%; Pred. No. 9.7e-61;
Matches 183; Conservative 88; Mismatches 231; Indels 53; Gaps 11;
QY 1 MRDYDEVIAFLGEMGFQRLIFELL---SASTIPNGFMGMSVVFVLAGTPHERCRVDPDAAN 57
Db 1 mptvddvleqvgefgvqkqafillclislap---lyvgivlfgftpgygcqpgvae 57
QY 58 LSS--AWR-----NNSVP-LRLRDGVEPHSCSYRL-----ATIANPSAL 95
Db 58 lsqrcgwsqaeeelnlyvpvgpsdeasflsqcmryevdngstldcvdplslvnsrql 117
QY 96 GLEPGRDVDLGLEQESCLDGEFSDQVYVLTSTVTEWNLVCEDNKNKVLPTTSLFFVGVLL 155
Db 118 plgp-----cehgwy--dtpgssivtefnlvcgdawkvdlfgscvnlgffl 162
QY 156 GSFVSGQLSDRGRKNVLPATMAVOTGFSFLQIFISWEMFTVLVIVMGQISNYVAF 215
Db 163 gslvvgviadrfgrklcllvtlvtvsvglvtavapdytsmlilrlldgmvsksqswsgy 222
QY 216 ILGTEILGKSVR----IIFSTLGCTFFAVGYMLPLFPAYFIRDRWMLLALTVPGLCV 271
Db 223 tlitefvsgyrrttailqma-----flvglvglagvayalpdrwrlqlavslptfifi 277
QY 272 PLWVFIPESRWLISQRRFREAEDIIQAAKNMNTAVPAVIFDSVEELNPLKQAKAFILD 331
Db 278 lywvfypesprwlisqkrttravimeqlaekngkvppadklmclcedasekrspsad 337
QY 332 LFRTRNIAIWTMSLLWMLTSVGYFALSIDAPNLHGDAYLNCFLSALIEIPAYITAWLL 391
Db 338 lfrtptlrkhtvilmlywfscavlygglhmhvgatganlyldffysvvefpaafiiivt 397
QY 392 LRLTPRRYIIAAVLEFWGGVLLFIQOLVPDYYFLSLGLVLMGLKFGITSFSLMYVFTAE 451
Db 398 idrigriyplaaenlvtgaacilmifiphelhwlntvltacigrmgatvlgmncvlnael 457
QY 452 YPTLVRNMAVGVTSTASRVGSIIPYFYV-LGAYNRMLPYIVNGSITVLIGITFTFFPES 510
Db 458 yptfirlngmmvcsalcdlggiftfmvrlmevqalplilgvlgtagantlllpet 517
QY 511 LGMTLPETLEOMQKV 525
Db 518 kgvalpetieeaenl 532

```

Search completed: January 25, 2002, 10:09:12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:08:24 ; Search time 23.33 Seconds
(without alignments)
531.475 Million cell updates/sec

Title: US-09-521-195-1
Perfect score: 2845
Sequence: 1 MRDDEVIAFLGEGWPFQRL.....KKTRDSMETEENPKVLITAF 551

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	702	24.7	555	US-08-501-572-3	Sequence 3, Appli
2	702	24.7	555	US-09-040-444-3	Sequence 3, Appli
3	687.5	24.2	556	US-08-501-572-1	Sequence 1, Appli
4	687.5	24.2	556	US-09-040-444-1	Sequence 1, Appli
5	680	23.9	553	US-08-501-572-2	Sequence 2, Appli
6	680	23.9	553	US-09-040-444-2	Sequence 2, Appli
7	630	22.1	537	US-08-647-397-2	Sequence 2, Appli
8	267	9.4	520	US-08-964-127-2	Sequence 2, Appli
9	234.5	8.2	494	US-09-031-392-5	Sequence 5, Appli
10	234.5	8.2	494	US-09-299-549-5	Sequence 5, Appli
11	222	7.8	492	US-08-355-844-3	Sequence 3, Appli
12	222	7.8	492	PCT-US95-16126-3	Sequence 3, Appli
13	216	7.6	493	US-09-031-392-10	Sequence 10, Appl
14	216	7.6	493	US-09-299-549-10	Sequence 10, Appl
15	213.5	7.5	524	US-08-928-692-12	Sequence 12, Appl
16	203	7.1	286	US-08-964-127-4	Sequence 4, Appli
17	196	6.9	500	US-09-031-392-7	Sequence 7, Appli
18	196	6.9	500	US-09-299-549-7	Sequence 7, Appli
19	181	6.4	509	US-09-031-392-6	Sequence 6, Appli
20	181	6.4	509	US-09-299-549-6	Sequence 6, Appli
21	174.5	6.1	584	US-08-928-692-13	Sequence 13, Appl
22	172	6.0	383	US-09-031-392-3	Sequence 3, Appli
23	172	6.0	383	US-09-299-549-3	Sequence 3, Appli
24	168.5	6.0	488	US-08-928-692-11	Sequence 11, Appl
25	168.5	5.9	488	US-08-928-692-10	Sequence 10, Appl
26	162	5.7	563	US-09-031-392-2	Sequence 2, Appli
27	162	5.7	563	US-09-299-549-2	Sequence 2, Appli

28	158.5	5.6	109	2	US-08-647-397-4	Sequence 4, Appli
29	156.5	5.5	534	2	US-09-031-392-4	Sequence 4, Appli
30	156.5	5.5	534	4	US-09-299-549-4	Sequence 4, Appli
31	149.5	5.3	322	4	US-08-964-127-6	Sequence 6, Appli
32	110	3.9	473	1	US-08-597-236-13	Sequence 13, Appli
33	110	3.9	473	1	US-08-746-682A-13	Sequence 13, Appli
34	110	3.9	548	3	US-08-903-139B-9	Sequence 9, Appli
35	109.5	3.8	834	2	US-08-677-734A-9	Sequence 9, Appli
36	109.5	3.8	834	2	US-08-677-734A-10	Sequence 10, Appli
37	108.5	3.8	1299	4	US-08-460-900C-62	Sequence 62, Appli
38	108.5	3.8	1299	4	US-08-674-509B-48	Sequence 48, Appli
39	108.5	3.8	1299	4	US-08-954-698-48	Sequence 48, Appli
40	107	3.8	1285	2	US-08-540-406-6	Sequence 6, Appli
41	107	3.8	1285	2	US-08-656-055-6	Sequence 6, Appli
42	107	3.8	1285	4	US-08-954-668-6	Sequence 6, Appli
43	107	3.8	1285	5	PCT-US95-13233-6	Sequence 6, Appli
44	103.5	3.6	1286	4	US-09-268-140-3	Sequence 3, Appli
45	103	3.6	548	3	US-08-903-139B-28	Sequence 28, Appli

ALIGNMENTS

RESULT 1
US-08-501-572-3
; Sequence 3, Application US/08501572
; Patent No. 6063623
; GENERAL INFORMATION:
; APPLICANT: Koepsell, Hermann
; APPLICANT: Grundeman, Dirk
; APPLICANT: Gorboulev, Valentin
; TITLE OF INVENTION: Transport protein Which Effects The
; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,
; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan,Henderson,Farabow,Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/501,572
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Toohey, Kimberlin M
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 02481.1453-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 555 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-501-572-3

Query Match 24.7%; Score 702; DB 3; Length 555;
Best Local Similarity 33.7%; Pred. No. 5.3e-65;
Matches 187; Conservative 93; Mismatches 229; Indels 46; Gaps 14;
QY 5 DEVIAFLGEGWPFQRLIFF---LLSASIIIPNGMNSVYFLAGTPEHRCRVPDAANLS-- 59

Db	6	DDVLEHGGEFFHQXQWFFALLSATP---	IYGVIFLGTFPHRCSRPGVAELSR	62
QY	60	SAWR-----NNSVPLRUROGREVPHSCSYRLA-TIANFSALGEPGRDVLQGLQESQ	113	
Db	63	COWSPAELNVTYFPGPAGEASPROCRYEYEDMNQSTFDCVPLASLDTNRSRLPLGPC	122	
QY	114	LDGWEFFSODVYLSVTFENMLVCEDNNKVPJLTTSLFFYGVLLGSFVGQSLSDRFRGNVL	173	
Db	123	RDGWY--ETRGSIVTEFNLVCANSKMWLDLFQSSVNVVFFIGSMSIGYIADRGKCLC	180	
QY	174	FATMAVQTGSFQLFESISHEMFTVLFVIVMGQISNYVAFILCTILKSKSVRIIFSL	233	
Db	181	LTTVLINAAAGVLMAISPTTMMJLPIRLTOGLVSKAGMLIGYILITEFVGGRV---	237	
QY	234	GVC--TFFANGYMLPLFAYFIRDNRMLLLATVPGVLCVPLMWFIPESRMLISORRR	291	
Db	238	GIFYOVAYTVGLLVLAGVAVALPHWRWLOFTVALPNFFFLLYWCIPESRMLISQKNA	297	
QY	292	EAEEDIIOKAAKMNNTVPAYI-----PDSVEELNPLAQOKAFITDLFRTRNTAI	345	
Db	298	EAMRIIKHAKNGKSLPASLOJRLREBETGKLAN-----SFLDLVRTPQIKRHTML	351	
QY	346	LLLMWLTVSGVFYALSOPAKHGD-NYLNCFYSALIEIPAYITAMLLRLTLPRYITAAV	404	
Db	352	MYNWFPTSVLVQGLIMHM-GLAGDNTYLDFFYSALVEEPAAFMILIIDRIGRYPWAA	410	
QY	405	LPWGGGVLLFTQLVPVDVYFELSGLVMGLKFGITSAFSMLYVFTAEILYPTLVNRNMA	464	
Db	411	NNVAGAACLASVFTFGDLQWLKIIISCLRGMGITWAYEIVCLVNAELYPTFI	470	
QY	465	STASRVGSIIPYFVILGAYNRM-----LPVIMGSLTVLIGITLFFPESIGMTLPET	518	
Db	471	SNMCDIGIITFFLVY-----RUTNLWEJLPMVFGVLVAGGIALCLPETA	525	
QY	519	LEQMOKYKWFRRSGKK	533	
Db	526	IEEAENMQRPNKKE	540	

RESULT 2

US-09-040-444-3
Sequence 3, Application US/09040444
Patent No. 6063766
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Grundeman, Dirk
APPLICANT: Gorboulev, Valentin
TITLE OF INVENTION: Transport protein which Effects The
TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,
TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,444
FILING DATE: March 18, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Connor, Steven P
REGISTRATION NUMBER: 41,225
REFERENCE/DOCKET NUMBER: 2481.1453-01
TELECOMMUNICATION INFORMATION:

STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/501,572
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Toohy, Kimberlin M
 REGISTRATION NUMBER: 35,391
 REFERENCE/DOCKET NUMBER: 02481.1453-00000
 TELEPHONE: (202)408-4000
 TELEFAX: (202)408-4400
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 556 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-501-572-1

Query Match 24.2%; Score 687.5; DB 3; Length 556;
 Best Local Similarity 33.2%; Pred. No. 1.8e-63;
 Matches 184; Conservative 88; Mismatches 230; Indels 53; Gaps 11;

QY 1 MRDYDEVIAFLGEGWGFQRLIFLL---SASIPNGFNMGMSVFLAGTPEHRCRVPDAA 57
 DB 1 MPTVDDVLEQGEFGWFGKQAFLLCLLSASLAP---IYVGIVFLGFTPGHYCONPGVAE 57

QY 58 LSS--AWR-----NNSVP-LRLRDGREVPHSCSYRL-----ATTANFSAL 95
 DB 58 LSQRCGWSQAEELNYVPLGLPSDEASFLSOCMRVEVDNQSTLDCVPLSLVANRSQ 117

QY 96 GLEPGRDVLGQLEQSCLDGWEFSQDVLSTVVTENWLVCEENKVKPLTTSLFFVGVLL 155
 DB 118 PLGP-----CEHGMY--DTPGSSIVTEFNLCGDAWKVDLFCVNLGFFL 162

QY 156 GSFVSGQLSDRGRKNVLFATMAVOTGFSFLOIFSINEMFTVLVIVGMQGISNYYVAF 215
 DB 163 GSLVGVYIADREFGRKLCILVTLVTSVSGVLTAVAPDYSMLLFFLQGMVSKGSWSGY 222

QY 216 ILGTIELGKSVR---IIFSTLGVCTFFAVGYMLPLPAYFIRDRWMLLALTVPGLCV 271
 DB 223 TLITEFVSGYRRTALLYQMA-----FTVGLVGLAGVAYAIIPDRWLQAVSLPTFL 277

QY 272 PLWVFIPESPRWLISORPREAEDIIOKAAMNNTAVPAVIFDSVEELNPLKQKAFILD 331
 DB 278 LYYWFVPSRWLLSOKRTRAVRIMEQIAKNGKVPVPPADLKMCLCEDEASEKRSPSFAD 337

RESULT 4
 US-09-040-444-1
 Sequence 1, Application US/09040444
 Patent No. 6063766
 GENERAL INFORMATION:
 APPLICANT: Koepsell, Hermann
 APPLICANT: Grundeman, Dirk
 APPLICANT: Gorboulev, Valentin
 TITLE OF INVENTION: Transport protein Which Effects The
 Transport Of Cationic Xenobiotics and/or Pharmaceuticals,
 TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow Garrett & Dunner, L.L.P.
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/040,444
 FILING DATE: March 18, 1998
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: O'Connor, Steven P
 REGISTRATION NUMBER: 41,225
 REFERENCE/DOCKET NUMBER: 2481.1453-01
 TELEPHONE: (202)408-4000
 TELEFAX: (202)408-4400
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 556 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-09-040-444-1

Query Match 24.2%; Score 687.5; DB 3; Length 556;
 Best Local Similarity 33.2%; Pred. No. 1.8e-63;
 Matches 184; Conservative 88; Mismatches 230; Indels 53; Gaps 11;

QY 1 MRDYDEVIAFLGEGWGFQRLIFLL---SASIPNGFNMGMSVFLAGTPEHRCRVPDAA 57
 DB 1 MPTVDDVLEQGEFGWFGKQAFLLCLLSASLAP---IYVGIVFLGFTPGHYCONPGVAE 57

QY 58 LSS--AWR-----NNSVP-LRLRDGREVPHSCSYRL-----ATTANFSAL 95
 DB 58 LSQRCGWSQAEELNYVPLGLPSDEASFLSOCMRVEVDNQSTLDCVPLSLVANRSQ 117

QY 96 GLEPGRDVLGQLEQSCLDGWEFSQDVLSTVVTENWLVCEENKVKPLTTSLFFVGVLL 155
 DB 118 PLGP-----CEHGMY--DTPGSSIVTEFNLCGDAWKVDLFCVNLGFFL 162

QY 156 GSFVSGQLSDRGRKNVLFATMAVOTGFSFLOIFSINEMFTVLVIVGMQGISNYYVAF 215
 DB 163 GSLVGVYIADREFGRKLCILVTLVTSVSGVLTAVAPDYSMLLFFLQGMVSKGSWSGY 222

QY 216 ILGTIELGKSVR---IIFSTLGVCTFFAVGYMLPLPAYFIRDRWMLLALTVPGLCV 271
 DB 223 TLITEFVSGYRRTALLYQMA-----FTVGLVGLAGVAYAIIPDRWLQAVSLPTFL 277

QY 272 PLWVFIPESPRWLISORPREAEDIIOKAAMNNTAVPAVIFDSVEELNPLKQKAFILD 331
 DB 278 LYYWFVPSRWLLSOKRTRAVRIMEQIAKNGKVPVPPADLKMCLCEDEASEKRSPSFAD 337


```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-040-444-2

Query Match      23.9%; Score 680; DB 3; Length 553;
Best Local Similarity 33.4%; Pred. No. 1.1e-62;
Matches 192; Conservative 85; Mismatches 228; Indels 70; Gaps 15;

QY 1 MRDYDEVIAFGEWCFQF---LIFFLLSASIIIPNGFNGSVVFLAGTPEHRCRVPPDAAN 57
DB 1 MPTVDIILEQGESGWFQKFLIICLLSAAFAP---ICVGIVFLGTPDHCOSPGVAE 57
QY 58 LSS--AWR-----NNSVPLRLRDGREVPHSCSYR-----LATIA-NFSALG 96
DB 58 LSQCGWSPAELNVTVPGLPAGEAFGLQCRRYEVDMNQSALSCVDPPLASLATNRSHLP 117
QY 97 LEPGRDVLGQLEQESCLDGHEFSDQVYLSVTVTNWNLVCEDNKVKVPLTTSFLFFGVLLG 156
DB 118 LGP-----CQDGWYI---DTFGSSIVTEFNLVCAWSKLDLDFQSLNAGFFFG 162
QY 157 SFVSGQLSDRFGRKNVLPATMAVQTFGFLQIFSISWEMFTVLFVIVGMQISNYVVAFI 216
DB 163 SLGVGYFADRFGRKCLLGLTVLVNAVSGVLMAFSPNYHSMILLFRLQLGLVSKGNMAGYT 222
QY 217 LGTEILG-----KSVRIIPSTLGVCTFFAVGYMLLPFLPAYFIRDMILLALTVPGVLCVP 272
DB 223 LITEFVSGSRRTVAIMYQMA-----FTVLGLVALTGLAYALPHWRMLQVLAFLSLPFL 277
QY 273 LWFIPESPRWLLSRRPREAEDIQKAAKMNTAVPAVI-----FDSVEELNPKQOK 326
DB 278 YMCVPEPRWLLSQRNTEAKIMDHIAQKNGKLPPADKMLSEEDVTEKLS----- 332
QY 327 AFTLDLFRTRNTAINTIMSLMLMLTSVGYFALSLDAPNLHGDAYLNCFLSALIEIPAYI 386
DB 333 -SPADLFRPLRKRTFILMYLWFTDSVLYQGLIHLHGATSGNLYLDLFYLSALVEIPGAF 391
QY 387 TAWMLLRTLPRIYIIAANVFWGGVLLFTQLVPDYFYFLSIGLVMLGKFGITSFASMLYV 446
DB 392 IALITIDRVGRIPMAVSNLLAGAACLVTFISP-DLHWLNIIIMCVGRMGITIAQMCL 450
QY 447 FTAEPLTVLRNMAVGVTSTASRVSGSIIAPYFY-IGAYNRMLPIYVMGSLAVLIGIFTL 505
DB 451 VNALIPTFVRNRVWVCSSLDIGIITFFIVFRLEWVQALPILFAVLGLAAGVTL 510
QY 506 FFPESLGMTLPETLEOMQKVKFRSGKTRDSMET 540
DB 511 LLPETKDALPETMKDAENL-----GRKAKPKENT 540

RESULT 7
US-08-647-397-2
Sequence 2, Application US/08647397
Patent No. 5972702
GENERAL INFORMATION:
APPLICANT: Beiter, David R.
APPLICANT: Brady, Kevin P.
TITLE OF INVENTION: OSTEOCLAST TRANSPORTER
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/647,397
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gates, Edward R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: B0801/7048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 537 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-647-397-2

```

```

Query Match      22.1%; Score 630; DB 2; Length 537;
Best Local Similarity 30.0%; Pred. No. 1.8e-57;
Matches 165; Conservative 102; Mismatches 245; Indels 38; Gaps 12;

QY 4 YDEVIAFLEWGEWCFQRLIFLFLSASIIIPNGFNGSVVFLAGTPEHRCRVPPDAANLSSAWR 63
DB 3 FSETLDVSGMGFPQYLHVLTLLALPILGIANHNLQIIFTATTPDHCHCRPPPNASL-EPW- 60
QY 64 NNSVPLRLRDGREVPHSCSYR---RLATIANFALSALGLEPGRDVLGQLEOESCLDGWEFS 120
DB 61 --VLPLG-PNGK--PEKCLRFVHLPLNASLNDTQGATEP-----CLDGWIYN 102
QY 121 QDVYLSVTWTVTNWNLVCEDNKVKVPLTTSFLFFGVLLGSFVSGQLSDRFGRKNVLFATMAVQ 180
DB 103 ST--RDTIVTDEWLVCVGNKLEKMAQSVFMAGILVGGPFGELSDRFGRKPILTWSYLL 160
QY 181 TGFSEFLQIFSISWEMFTVLFVIVGMQISNYVVAFILGTILGKSVRIIFS-TLGVCTFF 239
DB 161 AASGSSAAFSLSLTYMIFFLFCGCSISGLSTIILNVWVPTSTRAISSTIGYC--Y 218
QY 240 AVGYMLLPFLPAYFIRDMRMLLALTVPGLCVPLWFWFIPESPRWLLISQRRPREAEDIQK 299
DB 219 TIGQFILPLCLAYAVPQWRNLQLSVAFFIFSLLSWVPEISIRWLVLSGKFSRAKTLQR 278
QY 300 AAKMN-----NTAVPAVIFDSVEELNPLKQKAFILDLFRTRNTAINTIMSLMLMLT 352
DB 279 VATENGKKEGEKLTVEELKFNLDITSAR-VKYGSLDFRVSILRRVTFCLSLAWFAT 337
QY 353 SVGYFALSLDAPNLHGDAYLNCFLSALIEIPAYITAMLLRTPRRYIIAANVFWGGVLL 412
DB 338 GFAYYSLAMGVVEEFGVNIYIILQIFGGVDIPAKFITILSISYLGRRITOGFELLILAGVAI 397
QY 413 LFIQLVPDYFYFLSIGLVMLGKFGITSFASMLYVFTAEPLTVLRNMAVGVTSTASRVGS 472
DB 398 LALIFVSSEMOLLRTALAVFGKGLSGFSCLFLYTSELYPTVLRQTGMGTSNIWARVGS 457
QY 473 ITAPYFVYLGAYNRMPLPIYVMGSLVILGIFTLFPPEPSLGMTLPETLEOMQKVKFRSGK 532
DB 458 MIAPLVKITGELQEPINPVIEWMTMLLGGSAFFLLETNLRPLPETIEDIQ--DWYOOTK 515
QY 533 KTRDSMETEE 542
DB 516 KTKQEPFAEK 525

```

```

RESULT 8
US-08-964-127-2
Sequence 2, Application US/08964127
Patent No. 6277565
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
TITLE OF INVENTION: MOLECULES

```

NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/964,127
FILING DATE: 06-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-964-127-2

Query Match 9.4%; Score 267; DB 4; Length 520;
Best Local Similarity 23.4%; Pred. No. 2e-19;
Matches 117; Conservative 88; Mismatches 203; Indels 92; Gaps 21;

QY 85 RLATIANFSAIGLRGRDVGDLQEQSCLDGHEFSDQVYLSI-VYENWLVCEDNWKP 143
DB 50 RVATSDSCSGFAPP-----DFNHCKDMYDGLPVLTTNAGQWDLVCDLQWQVI 101
QY 144 LTTSFFVGVLLGFSVGSQSDREGRKNVLFATMAV-----QTGFSFLQIFS 190
DB 102 LQQLILFGLFASGYLFLGYPADREGRRGIVLLTLGLVPCGVGGAAGSSTGYMALR--- 158
QY 191 ISWEMFTLVIVGQISNVVAFILGTEILGKSVRIIFSTLGVCTFFFAVGTMPLPFA 250
DB 159 -----FLGLFLLAGVD-----LGVLMLRLECDPTQLRVALAGELVGGGHEFLGL-A 207
QY 251 YFIRDRMLLALTPGVLCVPLMW--FIPESPRWLISORRFEAEIDIOKAAMN----- 304
DB 208 LVSKDWRFLQRMITAPCLFLFYGPGLFLESARWLIVKQIEAQSVLRILAERNRPHG 267
QY 305 ----NTANPAVIFDSYELNPLKQKAF-ILDLPRTNIAIMTMSLLWLT----- 353
DB 268 OMLGEEAQUAL-QDLENTCLPLATTSFASCLLNRYNI-----WKNLLILGTFNFAHAI 321
QY 354 -----VG-----YFALSIDAPNLHGDAYLNC-FLSALIEIPAYITAWLLRLTPRY 399
DB 322 RHCYQPVGGGSPDFFYLCSLA---SSTAALACVFLG--VTVDREGRCILLSTWLAG 376
QY 400 ITAAVLFWGGGVLLFLQLVPVDYFISGLVLMGKFGITSAFMSLYVFAELYPPTLVNRM 459
DB 377 TASLVLL---GLWYDYNAAITTP-----SVLGLFSSQAAAILSTLLAAEVIPTTVRGR 427
QY 460 AVGTSTASRGSITAP-YFVYLGAYNRMLPYIVMGSULTVLIGIFLFFPESLGMTLPET 518
DB 428 GGLGIMALGALGSLGPAQRHMG-HGAFLQHVVLACALLCILSINMLPTEKRKLLPEV 486

QY 519 L---EQMKYKWPERSGKKTR 535
DB 487 LRDGELCRPSSLRLQPPPTR 506

RESULT 9
US-09-031-392-5
Sequence 5, Application US/09031392
Patent No. 5942398
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES
ENCODING GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-392-5

Query Match 8.2%; Score 234.5; DB 2; Length 494;
Best Local Similarity 25.2%; Pred. No. 4.8e-16;
Matches 115; Conservative 76; Mismatches 156; Indels 109; Gaps 23;

QY 133 NLVCEDNWKP-----LTTSFFVGVLLGFSVGSQSDREGRKN--VLFATMAV 179
DB 43 NYTLERSETPPSSVLLTSLWSLSVAIFSVGGMIGFSVGLFVNREGRRNSMLIVNLLAI 102
QY 180 QTG--FSLQIFSTISWEMFTVLFVIVGM--GQISNVVAFILGTEILGKSVRIEST--- 232
DB 103 AGGLMGECKI-AESVEMILGLRIILGFCGLCTGFVPMVI--GEISPTALGAGTLMQ 159
QY 233 LGVCTFTFANGYMLPLPAYFI----RDWRMLLALAV-PGVLCVPLWFWIPESPRMLISQ 287
DB 160 LGI---VIGILVAQIFGLKVLGTEDLWPLLLGFTILPAIIQCAALPFCSPESPRFLIN 215
QY 288 RFRFE-----AEDIIQKAAMNNTAVPAVIFDSVEELNPLKQKAFITLDFRTR 336
DB 216 RKEEKAKELQRLWNGFEDVAQIQENKD-----ESMRMSQEKQVTVLELFRAP 264
QY 337 N-----IAIMTMSLLWLTSGVYFALSIDAPNLHGDAYLNCFLSALIEIPAYIT--- 387
DB 265 NYRQPIIISIMQLSQSLGSGINAVFYSTGI-----PKDAGVQEPVATIGA 311
QY 388 -----AWLLRLTPRYITIAAVLFWGG-----GVLLFTQLVPVDYI-----ELSIG- 428

Db 312 GVNNTFTVSVFVVERAGRRL--HLIGLGMFAFCSILMTISLLLLKDNYSWMSFICIGA 369
QY 429 -LVMLGKGITSAFMSLVFVETAEYPTLVNMAVGVSTAS-----RVGSIAPYFVYLG 482
Db 370 ILVFAVFEIGPG-PIPFVFAELFGQGPRAAMAVAGCSNWTNELVGLLFPSSATEYLG 428
QY 483 AYNRMPLPIVMSLTVLIGITLFP-PPESLGMTLPE 517
Db 429 AY----VFIVTVFLVIFWVTFEFPKVPETGRTFEE 460

RESULT 10
US-09-299-549-5
; Sequence 5, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-549-5

Query Match 8.2%; Score 234.5; DB 4; Length 494;
Best Local Similarity 25.2%; Pred. No. 4.8e-16;
Matches 115; Conservative 76; Mismatches 156; Indels 109; Gaps 23;

QY 133 NLVCDNKKVP-----LTSLFFVGVILGSEVSGQLSDRFRGKN--VLFAFMAY 179
Db 43 NYLSESETPPSSVLTLSLSLVAIFSVMGIGSFVGLFVNFRGRNRMLIVNLLAI 102
QY 180 QTG--FSPQLQTSWEMFTVLFVIVGM--GQISNYVAVILGTEILGKSVRIEST--- 232
Db 103 AGGCLMGFKI-ABSVEMLILGRLLIIGLCGLCTGTFVPMYI--GETSPALRGAFGLNQ 159
QY 233 LGVCTTFVAVGYMLPLPAYFI-----RWRMLLALTVPGLVCLVPLWVFIPESPRWLIQS 287
Db 160 LGI-----VIGILVAVQIFGLKVLIGTEDLWPLLGLFTILPAIQCAALPFCPESPRFLIN 215
QY 288 RRFRE-----AEDIIQKAAMNNTAVPAVIFDSVEELNPLKQOKAFILDLPRT 336

Db 216 RKEERAKELLQRLWGTEDVAQDIQEMKD-----ESMRMSQEKQVTVLELFRAP 264
QY 337 N-----IAIMTMSLLMLLWLTSGVGFALSLDAPNLUHGDAYLNCFLSALIEIPAYIT--- 387
Db 265 NYRQPIIISINLISQOLSGINAVFYSTGI-----FKDAGVQVEFVATICA 311
QY 388 -----AWLLRLTPRYIIAAVLEWGG-----GVLLFIQLVDPVYI-----FLSIG- 428
Db 312 GVNNTFTVSVFVVERAGRRL--HLIGLGMFAFCSILMTISLLLLKDNYSWMSFICIGA 369
QY 429 -LVMLGKGITSAFMSLVFVETAEYPTLVNMAVGVSTAS-----RVGSIAPYFVYLG 482
Db 370 ILVFAVFEIGPG-PIPFVFAELFGQGPRAAMAVAGCSNWTNELVGLLFPSSATEYLG 428
QY 483 AYNRMPLPIVMSLTVLIGITLFP-PPESLGMTLPE 517
Db 429 AY----VFIVTVFLVIFWVTFEFPKVPETGRTFEE 460

RESULT 11
US-08-355-844-3
; Sequence 3, Application US/08355844
; Patent No. 5940307
; GENERAL INFORMATION:
; APPLICANT: Fischbarg, Jorge
; APPLICANT: Czegledy, Ferenc
; APPLICANT: Iserovich, Pavel
; APPLICANT: Li, Jun
; APPLICANT: Cheung, Min
; TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
; TITLE OF INVENTION: STRUCTURE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/355,844
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Tang, Henry Y.S.
; REGISTRATION NUMBER: 29,705
; REFERENCE/DOCKET NUMBER: A29927-50/29910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-408-2586
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 492 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..492
; OTHER INFORMATION: Facilitative glucose transportor
; OTHER INFORMATION: Glut1 protein
US-08-355-844-3

```
Query Match
Best Local Similarity 7.8%; Score 222; DB 2; Length 492;
Matches 117; Conservative 74; Mismatches 175; Indels 122; Gaps 21;

QY 119 FSDQVYLSTVTEWNLVCEENKVKPLTSLFFGVVLGSLGVSQSLSDRGRK-----NV 172
DB 52 YGESLPTTLTTLWS-----LSVAIFSVGGMISGFSVGLFVFNFRGRNSMLMNL 101
QY 173 LFATMAVQTGFSFLQIFSIWEM-----FTVLVIVGNGOISNVVAFILGT 219
DB 102 LAFSAVLMGSKL---GKSEFEMILGRFIIGVYCGLTGTFVPMYVGEVSPFAFRGALGT 158
QY 220 EILGKSVRIESTLGVCTFFAVGYMLPLPAYFI-----RD-WRMLLALTVPGLVCLVPLM 274
DB 159 VTILEFRSPAYRQPIILIAVVLQSLGGINAVFYTSI-----FEKAGV 303
QY 381 EIPAVIT-----AWLLRLTPRR--YIIAAVLFWGGVLLFIQLVPVDYV-- 423
DB 304 QQPVTATIGSGIVNTATVVSILFVVERAGRTHLIGLAGMAGQAILMTIALLEQLPW 363
QY 424 --FLSIGVLMGKFGITSAFS-----MLYVFTAELPTLVNMAVGVSTASRVGS-IIA 475
DB 364 MSYLSIVAI---FGVAFFEVGPGPIPMFIVAELESQGPRAAIAVAGFSNWTSNFIVG 419
QY 476 PFYVILGAYNMLPYIVNGSLTVLIGITLTF--PPESIGMTLPTLEQMOKVKWFRSGKKT 534
DB 420 MCFQVVEQLCGPYVFIETVLLVLFRTYFKVPKGRTFDEI-----ASGFRQGAS 473
QY 535 RDSMETEE 542
DB 474 QSDKTPEE 481

RESULT 12
PCT-US95-16126-3
; Sequence 3, Application PC/TUS9516126
; GENERAL INFORMATION:
; APPLICANT: Fischbarg, Jorge
; APPLICANT: Czegledy, Ferenc
; APPLICANT: Iserovich, Pavel
; APPLICANT: Li, Jun
; APPLICANT: Cheung, Min
; TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16126
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/355,844
; FILING DATE: 14-DEC-1994
; ATTORNEY/AGENT INFORMATION:
```

```
NAME: Tang, Henry Y.S.
REGISTRATION NUMBER: 29,705
REFERENCE/DOCKET NUMBER: A29927-50/29910
TELEPHONE: 212-408-2586
TELEFAX: 212-765-2519
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 492 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Human
FEATURE:
NAME/KEY: Peptide
LOCATION: 1,492
OTHER INFORMATION: Facilitative glucose transporter
OTHER INFORMATION: Glut1 protein
PCT-US95-16126-3

Query Match
Best Local Similarity 7.8%; Score 222; DB 5; Length 492;
Matches 117; Conservative 74; Mismatches 175; Indels 122; Gaps 21;

QY 119 FSDQVYLSTVTEWNLVCEENKVKPLTSLFFGVVLGSLGVSQSLSDRGRK-----NV 172
DB 52 YGESLPTTLTTLWS-----LSVAIFSVGGMISGFSVGLFVFNFRGRNSMLMNL 101
QY 173 LFATMAVQTGFSFLQIFSIWEM-----FTVLVIVGNGOISNVVAFILGT 219
DB 102 LAFSAVLMGSKL---GKSEFEMILGRFIIGVYCGLTGTFVPMYVGEVSPFAFRGALGT 158
QY 220 EILGKSVRIESTLGVCTFFAVGYMLPLPAYFI-----RD-WRMLLALTVPGLVCLVPLM 274
DB 159 VTILEFRSPAYRQPIILIAVVLQSLGGINAVFYTSI-----FEKAGV 303
QY 381 EIPAVIT-----AWLLRLTPRR--YIIAAVLFWGGVLLFIQLVPVDYV-- 423
DB 304 QQPVTATIGSGIVNTATVVSILFVVERAGRTHLIGLAGMAGQAILMTIALLEQLPW 363
QY 424 --FLSIGVLMGKFGITSAFS-----MLYVFTAELPTLVNMAVGVSTASRVGS-IIA 475
DB 364 MSYLSIVAI---FGVAFFEVGPGPIPMFIVAELESQGPRAAIAVAGFSNWTSNFIVG 419
QY 476 PFYVILGAYNMLPYIVNGSLTVLIGITLTF--PPESIGMTLPTLEQMOKVKWFRSGKKT 534
DB 420 MCFQVVEQLCGPYVFIETVLLVLFRTYFKVPKGRTFDEI-----ASGFRQGAS 473
QY 535 RDSMETEE 542
DB 474 QSDKTPEE 481

RESULT 13
US-09-031-392-10
; Sequence 10, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
```

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-031-392-10

Query Match 7.6%; Score 216; DB 2; Length 493;
Best Local Similarity 24.2%; Pred. No. 4.le-14;
Matches 111; Conservative 78; Mismatches 171; Indels 98; Gaps 19;

QY 144 LTTSLFFVGLGSGVSGQLSDRGKKNVLA--TMAVOTGFSE--LQIFSIWEMFTVLF 200
DB 68 LSVSIFAVGGMIGSLVXIGNRGRKXAMLVNNVLAAGLLMGLAKAXSEFEMIILGR 127
QY 201 VIVGM-----GOISNVYVAFILGT-EILGKSVRIIFS-TLGVCTFFAVGYML 245
DB 128 FIILGCLSSGVVPMVGEISPTALRGALGTNLQIGVIGILIAQVLGLDSL--LGNES 185
QY 246 LPLFAYFIRDMRLALLTPGVLCVPLWFWFIPESRWLI-----SORREREAE 294
DB 186 L-----WPLLGLTGPALLQLLLPFCPEPRYLLINKNEEARAKKALQRLGTA 236
QY 295 DIIOKAAMNNTAVPAVIFDSVEELNPLKQKAFILDIFRTRN-----IAIMTMSLLL 348
DB 237 DVSQEAEMKD-----ESRXMXSEKXSVLELFRSRXYRQPVIIAIVLQLSQOL 285
QY 349 WMLTSVGYFALSADPNLHGDAYLNCFLSALIEIPAYIT-----AWLLLRTL 395
DB 286 SGINAVFYISTSI-----FEKAGVQPVYATIGAGVNVNFTVVSFVVERA 332
QY 396 PRP--YIIAAVLFWGGVLLFIQLVPDY--YFLSIGVLMGKFGITSASF-----MLYV 446
DB 333 GRRTLHLGLGGMAGCAVMTIALALDQVPWMSYVSIVAI--FGFVAFEEVGPPIPF 390
QY 447 FTAELYPTLVRNMAVGTSTASRVGS-IIAPYFVYLGAYNRMIPYIVMGSLVLIIGFT- 504
DB 391 IVAELFSQGRPRAAIAVAGFNSNFTNFVGLLFOYIAELLGPYVFIIVFAVLLLIFFITF 450
QY 505 LFFPESLGMTLPELEQMKQKWFERSGKKTRDSMETEE 542
DB 451 LKVPETKGTETDEIAAARFXN--KXEQPEKESIEELE 486

RESULT 14
US-09-299-549-10
; Sequence 10, Application US/09299549

Patent No. 6136547
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.
ADDRESSEE: Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-APR-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 493 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-10

Query Match 7.6%; Score 216; DB 4; Length 493;
Best Local Similarity 24.2%; Pred. No. 4.le-14;
Matches 111; Conservative 78; Mismatches 171; Indels 98; Gaps 19;

QY 144 LTTSLFFVGLGSGVSGQLSDRGKKNVLA--TMAVOTGFSE--LQIFSIWEMFTVLF 200
DB 68 LSVSIFAVGGMIGSLVXIGNRGRKXAMLVNNVLAAGLLMGLAKAXSEFEMIILGR 127
QY 201 VIVGM-----GOISNVYVAFILGT-EILGKSVRIIFS-TLGVCTFFAVGYML 245
DB 128 FIILGCLSSGVVPMVGEISPTALRGALGTNLQIGVIGILIAQVLGLDSL--LGNES 185
QY 246 LPLFAYFIRDMRLALLTPGVLCVPLWFWFIPESRWLI-----SORREREAE 294
DB 186 L-----WPLLGLTGPALLQLLLPFCPEPRYLLINKNEEARAKKALQRLGTA 236
QY 295 DIIOKAAMNNTAVPAVIFDSVEELNPLKQKAFILDIFRTRN-----IAIMTMSLLL 348
DB 237 DVSQEAEMKD-----ESRXMXSEKXSVLELFRSRXYRQPVIIAIVLQLSQOL 285
QY 349 WMLTSVGYFALSADPNLHGDAYLNCFLSALIEIPAYIT-----AWLLLRTL 395
DB 286 SGINAVFYISTSI-----FEKAGVQPVYATIGAGVNVNFTVVSFVVERA 332
QY 396 PRP--YIIAAVLFWGGVLLFIQLVPDY--YFLSIGVLMGKFGITSASF-----MLYV 446
DB 333 GRRTLHLGLGGMAGCAVMTIALALDQVPWMSYVSIVAI--FGFVAFEEVGPPIPF 390
QY 447 FTAELYPTLVRNMAVGTSTASRVGS-IIAPYFVYLGAYNRMIPYIVMGSLVLIIGFT- 504
DB 391 IVAELFSQGRPRAAIAVAGFNSNFTNFVGLLFOYIAELLGPYVFIIVFAVLLLIFFITF 450

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:08:24 ; Search time 28.31 seconds
(without alignments)
1482.592 Million cell updates/sec

Title: US-09-521-195-1
Perfect score: 2845
Sequence: 1 MRDYDEVIAFLGEGFPQRL.....KKTRDSMETEENPKVLITAF 551

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2218	78.0	557	2 JW0089	organic cation tra
2	2152	75.6	557	2 JE0346	high-affinity carn
3	754.5	26.5	576	2 T22509	hypothetical prote
4	699.5	24.6	593	2 JC4884	organic cation tra
5	681.5	24.0	556	2 S50862	organic cation tra
6	565	19.9	794	2 T27870	hypothetical prote
7	548	19.3	745	2 T16565	hypothetical prote
8	484.5	17.0	527	2 T01019	transport protein
9	475	16.7	539	2 C96758	probable protein
10	421.5	14.8	518	2 B86299	hypothetical prote
11	412.5	14.5	521	2 H86298	hypothetical prote
12	382.5	13.4	515	2 B96825	hypothetical prote
13	356	12.5	528	2 T21682	hypothetical prote
14	338	11.9	540	2 T25851	hypothetical prote
15	327	11.5	454	2 F73580	probable sugar tra
16	317	11.1	455	2 B83213	probable MFS trans
17	311	10.9	591	2 T30895	sugar transport pr
18	307	10.8	435	2 T15290	hypothetical prote
19	294.5	10.4	529	2 T23190	hypothetical prote
20	290	10.2	400	2 C69757	transporter homolo
21	286.5	10.1	478	2 T33985	hypothetical prote
22	281	9.9	422	2 G72234	hypothetical prote
23	276	9.7	461	2 D70073	metabolite transpo
24	275	9.7	520	2 T23545	hypothetical prote
25	274.5	9.6	443	2 E64725	yaaU protein - Esc
26	274.5	9.6	443	2 H85485	probable transport
27	272.5	9.6	442	2 A83122	probable MFS trans
28	271.5	9.5	524	2 T19030	hypothetical prote
29	271	9.5	423	2 T19030	hypothetical prote

ALIGNMENTS

RESULT 1

JW0089
organic cation transporter protein 2 - human
N;Alternate names: OCTN2
C;Species: Homo sapiens (man)
C;Date: 18-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 28-May-1999
C;Accession: JW0089
R;Wu, X.; Prasad, P.D.; Leibach, F.H.; Ganapathy, V.
Biochem. Biophys. Res. Commun. 246, 589-595, 1998
A;Title: cDNA sequence, transport function, and genomic organization of human OCTN2,
A;Reference number: JW0089; MUID:98289574
A;Accession: JW0089
A;Molecule type: mRNA
A;Residues: 1-557 <WUA>
A;Cross-references: GB:AF057164; MID:g3273740; PIDN:AAC24828.1; PID:g3273741
A;Experimental source: placenta
C;Comment: This transporter functions in the elimination of cationic drugs and other

Query Match 78.0%; Score 2218; DB 2: Length 557;
Best Local Similarity 75.9%; Pred. No. 1.6e-156;
Matches 423; Conservative 57; Mismatches 71; Indels 6; Gaps 3;

Qy	1	MRDYDEVIAFLGEGFPQRLIFLISASIIPIGNFGMSVVFAGTPEHRCRVPDAANLSS	60
Db	1	MRDYDEVIAFLGEGFPQRLIFLISASIIPIGNFGMSVVFAGTPEHRCRVPDAANLSS	60
Qy	61	AWRNNVPLRLRDGREVPHSCRYRLATIANFSALEPGRDVDLQLEQESCLDGWEFS	120
Db	61	AWRNNVPLRLRDGREVPHSCRYRLATIANFSALEPGRDVDLQLEQESCLDGWEFS	120
Qy	121	QDVYLSTVVTWNVLVCEDNWKVPLTTSLFFVGVLLGSFVSGQLSDRFGKKNVLEATMAVQ	180
Db	121	QDVYLSTVVTWNVLVCEDNWKVPLTTSLFFVGVLLGSFVSGQLSDRFGKKNVLEATMAVQ	180
Qy	181	TGFSFLQIFSIQSWEMFTLVFVIVGMGQISNVVAFILGTILGKSVRIIFSTGLGVCFFFA	240
Db	181	TGFSFLQIFSIQSWEMFTLVFVIVGMGQISNVVAFILGTILGKSVRIIFSTGLGVCFFFA	240
Qy	241	VGYMLLPFAFIRDRWMLLALIVPGVLCVPLWFWFIPESPRWLLISQRRFREAEDIQKA	300
Db	241	VGYMLLPFAFIRDRWMLLALIVPGVLCVPLWFWFIPESPRWLLISQRRFREAEDIQKA	300
Qy	301	AKMNNTPAVIFD--SVEELNPLKQKAFILDLPRTNIAIMTISMLLLWMLTSVGYFA	358
Db	301	AKMNNTPAVIFD--SVEELNPLKQKAFILDLPRTNIAIMTISMLLLWMLTSVGYFA	358
Qy	359	LSLDAPNLHGDAYLNCFLSALIEIPAYITAMLLLTLPRLRYIIAAVLPWGGLVLTOLV	418
Db	359	LSLDAPNLHGDAYLNCFLSALIEIPAYITAMLLLTLPRLRYIIAAVLPWGGLVLTOLV	418
Qy	419	PVDYFSLSLGLVNLKFGITSAFSLMYVFTAEYFTLVNMAVGVTSTASRGSIIAYPF	478
Db	419	PVDYFSLSLGLVNLKFGITSAFSLMYVFTAEYFTLVNMAVGVTSTASRGSIIAYPF	478

probable sugar tra
hypothetical prote
hypothetical prote
hypothetical prote
hexose transport p
synaptic vesicle p
hypothetical prote
glucose transport
hexose transport p
glucose transport
probable transport
glucose transport
hypothetical prote
transmembrane tran
glucose transport
glucose transport

Db 421 PPDIYLAIVLVYKRGVTAFAFVYVYTAELVPTVVRNMGVSVSTASRLGSLSPYF 480
 Qy 479 VYLGAYNRLPYIVVWGSITLVIGLITFTFFPESLGMTLPETLEOMOKVKWFRSGK---KTR 535
 Db 481 VYLGAIDRFLPYILMGSITLTALTFLPESGTPLDPTIDQMLRVKMGKHKRTPSHTR 540
 Qy 536 DSMETEENPKVL-ITAF 551
 Db 541 MKDQGERPTILKSTAF 557

RESULT 2
 JE0346
 high-affinity carnitine transporter, CT1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: JE0346
 R:Sekine, T.; Kusuhara, H.; Utsunomiya-Tate, N.; Tsuda, M.; Sugiyama, Y.; Kanai, Y.; Endo
 Biochem. Biophys. Res. Commun. 251, 586-591, 1998
 A:Title: Molecular cloning and characterization of high-affinity carnitine transporter
 A:Reference number: JE0346; MUID:99011422
 A:Accession: JE0346
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-557 <SEK>
 A:Cross-references: DDBJ:AB017260; NID:g3869208; PIDN:BAA34399.1; PID:g3869209

Query Match 75.6%; Score 2152; DB 2; Length 557;
 Best Local Similarity 73.1%; Pred. No. 1.2e-151;
 Matches 407; Conservative 68; Mismatches 76; Indels 6; Gaps 3;

Qy 1 MRDYDEVIATFLGEMGPFQRLIFLLSASIIIPNGFNMGSVVFLAGTPEHRCRVPDAAANLSS 60
 Db 1 MRDYDEVIATFLGEMGPFQRLIFLLSASIIIPNGFNMGSVVFLAGTPEHRCRVPDAAANLSS 60
 Qy 61 AWRNNSVPLRLDGRVPHSCSYRLATIANFSALGLEPGRVDLQLEQESCLDGEWFS 120
 Db 61 AWRNNSVPLRLDGRVPHSCSYRLATIANFSALGLEPGRVDLQLEQESCLDGEWFS 120
 Qy 121 QDYVLSVTVTENLVCEENKWKVPLTTSFFVGLGSGVSGQLSDRFGKKNVLFATMAVQ 180
 Db 121 KDVFSTIVTEWDLQCKDKWAPLITSLFFVGLMGSGFISGQSLDRFGKKNVLFATMAVQ 180
 Qy 181 TGSFLOIFSIWEMFTLVFVIGMQISNYVVAFTILGTEILGKSVRIITFTGLVCTFFA 240
 Db 181 TGSFLOIFSVNEMFTLVFVIGMQISNYVVAFTILGTEILGKSVRIITFTGLVCTFFA 240
 Qy 241 VGYMLPLFAYFTRDWRMLLALTVPGLVPLWVFIPESPRWLSQRFRFRAEDIIQKA 300
 Db 241 FGMVPLFAYFTRDWRMLLALTVPGLVPLWVFIPESPRWLSQRFRFRAEDIIQKA 300
 Qy 301 AKNNNTAVPAVIFD--SVBELNPLKQKAFIIDLFRTRNIAIMTMSLLMLTMSVGYFA 358
 Db 301 AKFNGIVASTIFDSELDQNSKRPQSHIYDLYVTRNIRIITMSLLMLTMSVGYFG 360
 Qy 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRYITAAVLFWGGGVLLFIOLV 418
 Db 361 LSLDTPNLHGDIVYVNCFLAAVEVPAYITAWLLQLHLPYISAAALFLGGSVLLFIOLV 420
 Qy 419 PVDYVFLSGLVLMGKFGTTSAFSMLYVFTAEYPTLVYRNMAVGTSTASRVGSI 478
 Db 421 PSELVFLSTALVWVGKFGTTSAYSVYVITAEYPTLVYRNMGVSVSTASRLGSLSPYF 480
 Qy 479 VYLGAYNRLPYIVVWGSITLVIGLITFTFFPESLGMTLPETLEOMOKVKWFRSGK---KTR 535
 Db 481 VYLGAIDRFLPYILMGSITLTALTFLPESGTPLDPTIDQMLRVKMGKHKRTPSHTR 540
 Qy 536 DSMETEENPKVL-ITAF 551
 Db 541 TOKDGGESPTVLKSTAF 557

RESULT 3
 T22509
 hypothetical protein F52F12.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22509
 R:Matthews, L.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19573
 A:Accession: T22509
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-576 <WIL>
 A:Cross-references: EMBL:Z83228; PIDN:CAB05732.1; GSPDB:GN00019; CESP:F52F12.1
 A:Experimental source: clone F52F12
 C:Genetics:
 A:Gene: CESP:F52F12.1
 A:Map position: 1
 A:Introns: 16/1; 23/3; 50/1; 80/3; 108/2; 134/3; 221/1; 257/2; 319/3; 358/2; 490/3; 5

Query Match 26.5%; Score 754.5; DB 2; Length 576;
 Best Local Similarity 30.2%; Pred. No. 2.8e-48;
 Matches 169; Conservative 119; Mismatches 236; Indels 35; Gaps 9;

Qy 3 DYDEVIAFLGEMGPFQRLIFLLS-ASIIIPNGFNMGSVVFLAGTPEHRCRVPDAAANLSSA 61
 Db 16 DFDFLEQVGNVGYQIVFFFIICLPTSLPSAFSAFNIPFVYGNPHTCHIEPGEKYLRLP 75
 Qy 62 WRNNSVPLRLDGRVPHSCSYRLATIANFSALGLEPGRVDLQLEQESCLDGEWFSQ 121
 Db 76 LTNDTQIL-----SCHQYNETQINVFRAFTSAP-VDTYSDRISLVPCQNGWDYDN 124
 Qy 122 DYVLSVTVTENLVCEENKWKVPLTTSFFVGLGSGVSGQLSDRFGKKNVLFATMAVQ 181
 Db 125 STYLSLVTENLVCEENKWKVPLTTSFFVGLGSGVSGQLSDRFGKKNVLFATMAVQ 184
 Qy 182 GFSFLOIFSIWEMFTLVFVIGMQISNYVVAFTILGTEILGKSVRIITFTGLVCTFFAV 241
 Db 185 VGTASSPAKDIETFIILRFTGLAFPAFQIPFICMEFMGNSGR-IFSGLMTSLFFGA 243
 Qy 242 GYMLPLFAYFTRDWRMLLALTVPGLVPLWVFIPESPRWLSQRFRFRAEDIIQKAA 301
 Db 244 AMALLGVVAMFIRWRQLTPECNAPAFYIIYFELPESPRWSVSGKWADAKKQLKKA 303
 Qy 302 KMN---NTAVPAVIFDSVEELNPLKQKAF-----ILDLFRTRNIAIMTMSLLMLTS 353
 Db 304 KMGKSNVDVDELV-DSMKNHQNAEEKETKRSHNVTDLEKTPNLRKTLIVTYIWMNA 362
 Qy 354 VGYFALSIDAPNLHGDAYLNCFLSALIEIPAYITAWLLRLTPRYITAAVLFWGGGVLL 413
 Db 363 IYVGLNLVNSLNPVDYVSWFIINGAVELPGYFVWPLQACAGRRWTLAATMIVCGIGCV 422
 Qy 414 FIQLVPDYVFLSGLVLMGKFGTTSAFSMLYVFTAEYPTLVYRNMAVGTSTASRVGSI 473
 Db 423 SAMPMDPGPWLVASAFSGFKGSGFAVIYIFAGELYPTVVYRAICMGSSNVASGGL 482
 Qy 474 IAPYVYVGLGAYNRLPYIVVWGSITLVIGLITFTFFPESLGMTLPETLEOMOKVKWFRSGK 533
 Db 483 LAPHIVNLGKIVLPLLLMLLMLALSAGILTFFLPTLGLAPLMTTIEDAENF-----GKK 537
 Qy 534 -----TRDSMETEENP 544
 Db 538 PEPDSGMFTQAACKRESQP 556

RESULT 4
 JC4884
 organic cation transporter protein 2 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
 C:Accession: JC4884

R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
 Biochem. Biophys. Res. Commun. 224, 500-507, 1996
 A:Title: cDNA cloning and functional expression of a novel rat kidney organic cation transporter
 A:Reference number: JC4884; MUID:96295517
 A:Accession: JC4884
 A:Molecule type: mRNA
 A:Residues: 1-593 <ORU>
 A:Cross-references: DBJ:D83044; NID:g1502282; PIDN:BAAL1754.1; PID:g1502282
 A:Experimental source: kidney
 C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 24.6%; Score 699.5; DB 2; Length 593;
 Best Local Similarity 34.0%; Pred. No. 3.4e-44;
 Matches 189; Conservative 87; Mismatches 243; Indels 37; Gaps 14;

QY 1 MRDYEVIATFGECPFORLIFFLLSASIPNGFNG--MSVVFAGTPEHRCRVPDAANL 58
 DB 1 MSTVDIIEHGEFHLFOKQTFLL--ALLSGATPIYGVIFLGFTHHCWSPGAKL 58
 QY 59 SS--AWR-----NNSVP--LRLRDGREVPHSCSYR-----LATIANFSALGLEPGRDV 103
 DB 59 SQRCSQAEEELNYTPVGLGPSDEASFLSCMRYEVDWQSTLDCVDPPLSLA-----A 112
 QY 104 DLGQLEQSCLDGFEFSQDYLSTVVTENLVCEENKVPITLTSILFFVGLLGSFVSQGL 163
 DB 113 DRNQLPLGFCERHGWYNTFG--SSIVTEFLVCAHSMMLDLFQSVVNVGFFIGAMMIGYL 170
 QY 164 SDRFGKKNVLFATMAVOTGFSLQIFSIWEMFTVLFVIVGMQISNVYVAFILGTEILG 223
 DB 171 ADRFGKCLLVITILINISGALMAISPNYAWMLVFRFQGLVSKAGWLGILYILITEFVG 230
 QY 224 KSVRIIFSTLGYC--TFPANGVYMLPLFAYFIRDRWMLLLALTVPGLVCLVPLWFIPEP 281
 DB 231 LGYR--RMVGICYQIAFTVGLLILAGVAYTIPNNRWLQFAVTFNFCFLLYFMCIPESP 287
 QY 282 RWLISORFREADIIOKAKNNNTAVPAVIFDSVEELNPLKQKAFILDLFRTNIAIM 341
 DB 288 RWLISONKIVKAMKIHKAKNGKSVPSVLQNLTPDEDAGKKLPSILDVLTPIRKH 347
 QY 342 TMSLLMLMTSVGYFALSALDAPNLHGD--AYLNCFLSALIEPAITAWLLRLTPRYI 400
 DB 348 TLILMYNFTSSVLYQGLIMHM--GLAGDNIYLDFFYSALVEPFAAFIILITIDRGRYP 406
 QY 401 IAAVLFWGGVLLFIQVPDYFYLSIGLVMLKFGITSAFSLMVFVFAELPTLVRNA 460
 DB 407 WAVSNVAGAACLASVFPIDDLQWLKITACLGRMGITMAYEMVCLVNAELPTIYRNLG 466
 QY 461 VGTSTASRVGSIAPFYV--LGAYNRMLPYIVMGSLTVLIGITLFFFPESGLMTLPETL 519
 DB 467 VLVCSMCDIGIITFELYRLTDIWMFPLVFAVGVGVAGALVLLLPETKGLPETI 526
 QY 520 EQMKVWFRSGKKTR 535
 DB 527 EDAENMQ--RPRKKER 540

RESULT 5
 S50862
 Organic cation transport protein OCT1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Nov-1999
 C:Accession: S50862; S78533; I58089
 R:Gruendemann, D.; Gorboulev, V.; Gambaryan, S.; Veyhl, M.; Koepsell, H.
 Nature 372, 549-552, 1994
 A:Title: Drug excretion mediated by a new prototype of polyspecific transporter.
 A:Reference number: I58089; MUID:95082907
 A:Accession: S50862
 A:Molecule type: mRNA
 A:Residues: 1-556 <ORU>
 A:Cross-references: EMBL:X78855; NID:g633621; PIDN:CAA55411.1; PID:g633622
 A:Experimental source: kidney
 R:Gorboulev, V.G.

submitted to the EMBL Data Library, January 1995

A:Reference number: S78533
 A:Accession: S78533
 A:Molecule type: mRNA
 A:Residues: 1-342, 'N', 344-556 <OR>
 A:Cross-references: EMBL:X78855; NID:g633621; PIDN:CAA55411.1; PID:g633622
 C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
 F:20-46/Domain: transmembrane #status predicted <TM1>
 F:154-171/Domain: transmembrane #status predicted <TM2>
 F:178-197/Domain: transmembrane #status predicted <TM3>
 F:243-260/Domain: transmembrane #status predicted <TM4>
 F:267-283/Domain: transmembrane #status predicted <TM5>
 F:350-366/Domain: transmembrane #status predicted <TM6>
 F:380-398/Domain: transmembrane #status predicted <TM7>
 F:406-425/Domain: transmembrane #status predicted <TM8>
 F:435-452/Domain: transmembrane #status predicted <TM9>
 F:469-485/Domain: transmembrane #status predicted <TM10>
 F:494-514/Domain: transmembrane #status predicted <TM11>
 F:71,97,113,432/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:286,292/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status
 F:296,343,550/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status

Query Match 24.0%; Score 681.5; DB 2; Length 556;

Best Local Similarity 33.0%; Pred. No. 6.8e-43;
 Matches 183; Conservative 88; Mismatches 231; Indels 53; Gaps 11;

QY 1 MRDYEVIATFGECPFORLIFFLL--SASITPNGFNGSVVFLAGTPEHRCRVPDAAN 57
 DB 1 MPTVDDVLEQVGEFGFQKQAFLLCLLISASLAP---IYGVIFLFTPGHVCQNGVAE 57
 QY 58 LSS--AWR-----NNSVP--LRLRDGREVPHSCSYRL-----ATIANFSAL 95
 DB 58 LSQRCSQAEEELNYTPVGLGPSDEASFLSCMRYEVDWQSTLDCVDPPLSLVANSOL 117
 QY 96 GLEPRDVLGQLEQSCLDGFEFSQDYLSTVVTENLVCEENKVPITLTSILFFVGLL 155
 DB 118 PLGP-----CEHGWY--DTPGSSIVTEFLVCGDAWKVDLFQSCVNLGFFL 162
 QY 156 GSFVSGQLSDRFGRKNVLFATMAVOTGFSLQIFSIWEMFTVLFVIVGMQISNVYVAF 215
 DB 163 GSVVGVYADRFGRKCLLVLTLSVSGVLTAVAPDYTSMLLFRLLQGMVSGWVSGY 222
 QY 216 ILGTEILGKSVR-----IIFSLGVTFFANGVYMLPLFAYFIRDRWMLLLALTVPGLV 271
 DB 223 TLITEFVSGYRRTAILYQMA-----FTVLGLAGVAYAIIPDRWLQAVSLPTEFL 277
 QY 272 PLWFIPEPSPMLISORFREADIIOKAKNNNTAVPAVIFDSVEELNPLKQKAFILD 331
 DB 278 LYWFVPEPSPMLISQKRTTRAVRIMEQIAOKNGKVPADLKMCLCEEDASEKRSFAD 337
 QY 332 LFRTRNIAIMTMSLLMLTSVGYFALSALDAPNLHGDAYLNCFLSALIEPAITAWLL 391
 DB 338 LFRPTLRKHVILMYLWFSCAVLYQGLIMHVGTAGANLYLDFEYSSIVEPFAAFIILVT 397
 QY 392 LRTPLPRVYIIAALFWGGVLLFIQVPDYFYLSIGLVMLKFGITSAFSLMVFVFAEL 451
 DB 398 IDRIGRIYPIAASNLVTGAACLLMIFIPHEHLNVLTLACLGRMGATIVLQMVCLVNAEL 457
 QY 452 YPTLVNRNAVGTSPASRVGSIAPFYV--LGAYNRMLPYIVMGSLTVLIGITLFFPES 510
 DB 458 YPTFIRNLGMVCSALCDLGGITFPFVFRMLMEVWQALPLILFGVLGTAGAMTLLPET 517
 QY 511 LGMTLPETLEQMKV 525
 DB 518 KGVALPETIEEAENL 532

RESULT 6
 T27870
 hypothetical protein ZK455.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T27870

R:white, S.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z20432

A:Accession: T27870

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-794 <WIL>

A:Cross-references: EMBL:Z66567; PIDN:CAA91492.2; GSPDB:GN00028; CESP:ZK455.8

A:Experimental source: clone ZK455

C:Genetics:

A:Gene: CESP:ZK455.8

A:Map position: X

A:Introns: 52/3; 119/2; 158/1; 207/3; 291/3; 335/2; 428/3; 610/2; 636/3; 733/3

Query Match 19.9% Score 565; DB 2: Length 794;

Best Local Similarity 26.4% Pred. No. 4.2e-34;

Matches 158; Conservative 121; Mismatches 232; Indels 88; Gaps 20;

Qy 3 DYDE-----VIAFLGEGWGPORLIFFLLSASIIIPNGFMGSMVFLAGTPEHRCRVP-----53

Db 155 DFEEDLIGLOIGCSYQWIIIVLIISVQVQPHAMFNLSVVMYQPDHWCXIPFNEE 214

Qy 54 ----DAANLSSAW--RNNSV--PLRLRDGVEPH--SCSRRLATTANFSAALGLEPGRD 102

Db 215 SFSAEGLTNTWQVLSNTIAPRTENKORNELHDOCHYE-----RDYVHKLSFWAQ 270

Qy 103 V-DLQGLQESCLDQWESQDVYLSVTVNVLNVCEDNWKVPLTTLFVFGVLLGSFVSG 161

Db 271 VKDMNATKINRCEWEYDYSMDRTIVTEWNRVCDNNSRAHVMSYSLGYLVGCFVGG 330

Qy 162 QLSDRFGRKNVLEATMAVQTGFSFLQIFSISWEMFTVL-FVIIVGMGOISNYVAVILGTE 220

Db 331 FIDRYGRKTAITGEGILTMFLGELLYSKEFLFVVRFLAATNEAAD-LAAVVICWE 389

Qy 221 ILGKSVRIIFSTGLVCTFFAVGYMLLFLFYFIRDMRLMLLALTLPVGLCVPLWPFIPES 280

Db 390 VTGTYRSIVGSL-IOAPWAGYAFALAIYLTAKSWTMIHLICVLLHIIISMLLYPLPS 448

Qy 281 PRLWISORRFEADIIIOKAKMNTAVPA--VIFDSVEELNPLK-QOKAFILDLFRTN 337

Db 449 PRLWLNKTKOAEKILREACHYKRLPSDLGLVLRHAEEKWKWKNEKPSYFHLFRSE 508

Qy 338 IAIMTISMLLWM-----LTSVGY-----356

Db 509 LFRNVVLFIVWVTKNIFONSREFOYSDNSCSGLLRNGYCIVRSVITRKVGRPLSEKFR 568

Qy 357 -----FALSIDAPNLKGDAYLNCFLSALIELPAYITAWLLRLPLPRYIIAALVFGG 411

Db 569 NOKLCFRVFD-----GNFFLNNAAGAIETPLVFCVFLLR-MGRKRSQMLVLF-GSGL 621

Qy 412 LLFTQLVPV--DYFLSLGLVMLGKFGITSFASMLXYFTAEIYPTLVNRMVAVGTSTASR 469

Db 622 FLTTSVVMVVRKQSTLALIFLMLLSKACIQGSFNILYFTSELNPTVVRNSAVGSISSW 681

Qy 470 VGSIIAPFVVLGAYNR-MLPYIVMGSLTVLIGITLFFPESGLMTLPET-LEOMQKVK 526

Db 682 MGAGASGIIALSDVWPLVPLWTFITACFSLLACGLVLLLPETOGCLPLDPIILDSVQNV 740

RESULT 7

Ti6565

hypothetical protein K05F1.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: Ti6565

R:Wohlmann, P.

submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K05F1.

A:Reference number: Z18537

A:Accession: Ti6565

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-745 <WOH>

A:Cross-references: EMBL:U29377; NID:g868173; PID:g868176; PIDN:AAA68713.1; CESP:K05F

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:K05F1.6

A:Introns: 70/2; 160/3; 205/3; 249/3; 293/2; 350/1; 386/3; 462/3; 546/2; 669/3; 680/3

Query Match 19.3% Score 548; DB 2: Length 745;

Best Local Similarity 27.8% Pred. No. 7e-33;

Matches 166; Conservative 112; Mismatches 233; Indels 86; Gaps 21;

Qy 3 DYDEVIAFLGEGWGPORLIFFLLSASIIIPNGFMGSMVFLAGTPEHRCRVP-----53

Db 112 DFEILNIIGGCRWQIWIIVLIALQOIPHAMFNLSVVMYQPDHWCXIPFNEE 171

Qy 54 -----DAANLSSAWRN-NSVPLRLDRGVEPHSC-----SRYLATIANFS 93

Db 172 ISSHGPLYMGVDDIKNISVFPNANS-----DGAVQDSCFYERSEERYKQLRMPLE 225

Qy 94 ALGLEPGRVDLQGLEQESCLDQWESQDVYLSVTVNVLNVCEDNWKVPLTTLFVFG 152

Db 226 TAMSEANKDV-----APKKKC-QAYHFEKDVVETIVTDENLVC-DSWFAKGAHMFYSIG 279

Qy 153 VLLGSFVSGQLSDRFGRKNVLEATMAVQTGFSFLQIFSISWEMFTVL-FVIIVGMGOISNYV 212

Db 280 YLLGCVLGGTASDKIGRKPTIIIGFGLSSMLGVLPFDNYIPMLLIRLULSAICNEAADL 339

Qy 213 VAFILGTEILGKSVRIIFSTGLVCTFFAVGYMLLFLFYFIRDMRLMLLALTLPVGLCV 272

Db 340 AAYTLCMEITGTYRAMVGSW-LQAPWALGYALLAIYLTAKSWKTIQVIAAGLHFMISII 398

Qy 273 LWNIPSPRWLISORFREAEADIIQKAAK-----MNTA-----VPA---VIFDSVEE 318

Db 399 FICISIPSPRWLMVQNVSEAEVIRKACREPPFPNMCTTSCGNLPSDLDELIVSHREK 458

Qy 319 LNPLOKQKAFILDTRTNIAIMTISLLWMLTSVGYFALS-----DAPN---LHSDA 370

Db 459 LNKNGKGIKGLDFTMKELRYRTISVCIVFMATLVYGLVMALSDQSPAGRTLTGYF 518

Qy 371 YLNCFLSALTEIPA-YITAWLLRLPLPRYIIAALVFGWGGVLLFTQLVPVD-YVFLSIG 428

Db 519 HLNGIAGAIETPLFACVWMW--QLGRKALMLTITSGLFTIIVAMLSVSGHYMLALA 576

Qy 429 LVMLGKFGITSAFSLMYVFTAEIYPTLVNRMVAVGTSTASRGVSIAPYFVILGAYNRL 488

Db 577 FMYGKIAGVQAFNLIYFTSELYPTVVRNTAVGVSMVAREGSLSYIALLS--NISL 634

Qy 489 PYIVMGSLTVLIGITLTF-----FPESLGMWLPETLEQ-----MOKVWFRSGKKT 534

Db 635 PIVPM----IFAVFSLFAGMLVFLPETSEKPLPETDDAINFLEPTKQFTIEST 687

RESULT 8

T01019

transport protein homolog YUP8H12R.2 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 22-Oct-1999

C:Accession: T01019

R:Theologis, A.; Vysotskaya, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; K

Definer, P.; Davis, R.W.

submitted to the EMBL Data Library, May 1998

A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.

A:Reference number: Z14227

A:Accession: T01019

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-527 <THE>

A:Cross-references: EMBL:AC002986; NID:g2494106; PID:g3152576; GSPDB:GN00059; ATSP:YU

A:Experimental source: cultivar Columbia

C:Genetics:

A:Gene: ATSP:YUP8H12R.2

A:Map position: 1

Query Match 17.0%; Score 484.5; DB 2; Length 527;
Best Local Similarity 27.0%; Pred. No. 2.3e-28;
Matches 150; Conservative 99; Mismatches 209; Indels 97; Gaps 17;
QY 2 RYDEVIA-AFLGEMGPFORLIFFLLSIIIPNGFMGSMVFLAGTPEHRCRVPDANLSS 60
DB 18 RSLDDTIESYIGSGWAQALVAFSGVFDAAQTFLSVFTDSEPTWHC----- 67
QY 61 AWRNNSVPLRLDREVPHSCRSRYLATIANFSAALGPEGRDVLGQEQESCLDGNFES 120
DB 68 -----TDSNSICHE-----SISNICIL-----PKTAWSDIS 94
QY 121 QDVYLSVTWENLVCEENKWKVPLTTSFFVGLGSEFVSQGLSD-RFGRKNVLFATMAV 179
DB 95 PHV---SWISEWGLQACAGSFYKGLPESFFVGCCLIGGLVLTADSSSLGRKNMFLSLV 151
QY 180 QTGFSLQIFSIWEMFVFLVIVGMGOISNVVAFILGTEILGKSVRIIFSTLGVCCTFF 239
DB 152 MAISTMLTVFSPNIWYAVLRFVNGFGRATIGTCALVLSLTELVGKKWR---GRVGIIMSFF 208
QY 240 A--VGMPLPLFAYIR--DWRMLLALTVPGL--CVPLWFIPEPSPRLISQRRFREAE 294
DB 209 GEMGLFSLPLMAYNWRSSWRILYAWTSPIIYCVLVRVFCVSPRLVFRGRREAI 268
QY 295 DIITQAAKNMNTAVP-----AVIFDSV---EELNPLKQKAFILDLFRTRNIAIMTMS 345
DB 269 SILKRVASIPSTDVSSGGAISNFSFSLPDEDEKP-----STNVNIFTMK 315
QY 346 LLL---WMLTSGV-----YFALSLDAPNLHGDAYLNCFLSALIEIPAVITAWL 390
DB 316 VLVEKRWALKRLSAVMAIAFGICLVYIGPLALSNDLFNLYLSAFAFALMDLPANLITLF 375
QY 391 LRLTLPRRYIIAALFWGG--GVLLFI--OLVPVDYVYFSLGLVLMGLFGIITSFASFLYV 446
DB 376 LVDKLSRRNALGFTALGGVSSVLIFALHNMRIHNGALQALALELISYFSACSAFNMEMI 435
QY 447 FTAELTYTLVRNNAVGTSTASRVGSIIPYFYLYGAYNMLPYIVMGLSLTVLIGIFTLF 506
DB 436 YTIELEPTCVRSIAIARQALVLGVFSPIMVAAGRNKNAFWSFGIFGLAIGLLGLFAVG 495
QY 507 FPESLGLTLPETLEQ 521
DB 496 LPETRGSLDCLDTMDE 510

RESULT 9
C96758
probable protein transporter T18K17.11 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96758
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96758
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <STO>
A:Cross-references: GB:AE005173; NID:g6598860; PIDN:AAF18714.1; GSPDB:GN00141
A:Gene: T18K17.11
A:Map position: 1

Query Match 16.7%; Score 475; DB 2; Length 539;
Best Local Similarity 29.8%; Pred. No. 1.2e-27;
Matches 145; Conservative 82; Mismatches 193; Indels 66; Gaps 13;
QY 87 ATTANFSALGPEGRDVLGQEQESCLDGEF--SODVYLSVTWENLVCEENKWKVP 143
DB 97 AIVEGASLCLGASGE-----WEWIGPKSD-----TVVSEWNLICQHKFLVA 137
QY 144 LTTSLFFVGLGSEFVSQGLSDR-FGRKNVLFATMAVQTGFSLQIFSIWSE----- 194
DB 138 VPSTLFFIGSLFGSGVGYLADSWFGKRTLLLSCLV-----TFVTAFATSPNVWVYAF 193
QY 195 -MTVTLFVIVGMGOISNVVAFILGTEILGKSVRIIFSTLGVCCTFEAVGMPLPLFAYFI 253
DB 194 LRFANGFFRSIGGS-----CCIVLATEIVGKWRGQGVQYGF-FFTLGFLSLPLMAYLE 247
QY 254 R-DWRMLLALT-VPGVLCVPLWFIPEPSPRLISQRRFREAEIIQAAKNMNTAVPAV 311
DB 248 RKSRLNLYRIISPLPLGYAVCLLPAYESPRLLVGRNKEAMVLLKRLARLNGKQLPAD 307
QY 312 IFDSVEELNPLKQKAFILDLFRTRNIAIMTMSLLWMLTSGVYFALSIDAPNLHGDAY 371
DB 308 L-SLVDPIPERDDQTSSEKFKTKWAKRIIMVMAGFGSGFVYGIQNAENLNFLY 366
QY 372 LNCFLSALIEIPAVITAWLRLTLPRYIIAALVFWGGVLLFTQLVPVDY----- 423
DB 367 LTVAVNALMEFFPAVFIGSFLGLGVNRRPLFSNSSLAGFACILCAVLISHRVIRATSVAK 426
QY 424 FLISGLVMLKFGITSAFSMLYVFTAEYPTLVNRMNAVGTSTASRVGSIIPYFYVYVLA 483
DB 427 WLQAVEAVGFMASTAYDVLYVYVVELEPTNVNRTAVSLLRQAFMLGSAAPLLVALGR 486
QY 484 YNRMPLPYVMGLSLTVLIGIFTLFPFESLGMPLPTLQOMQKVKWFSKKTROSMETEEN 543
DB 487 ESAMMSFIVFGVASVLGSGIVSLRLRNNAPLYETLAQOGKA-----EEIEN 533
QY 544 PKVLIT 549
DB 534 ETIMIT 539

RESULT 10
B86299
hypothetical protein AAD34691.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B86299
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: B86299
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <STO>
A:Cross-references: GB:AE005172; NID:g4966360; PIDN:AAD34691.1; GSPDB:GN00141
A:Gene: T18K17.11
A:Map position: 1

Query Match 14.8%; Score 421.5; DB 2; Length 518;
Best Local Similarity 25.5%; Pred. No. 1e-23;
Matches 138; Conservative 90; Mismatches 233; Indels 81; Gaps 14;
A:Map position: 1

Db 98 WEWDGSGKSVISFGLGECSSLLRMPSPSAPYIGAIVGGFFLIALIPDDSLGRKKLVLP 157

Qy 175 ATMAVQTGSFLQIFSIWEMFTVLFVIGMGQISNVVAFTLGTETILGKSVRITESTLG 234
: : | : : : : | : | : : : : | : : : : | : : | : :

Db 158 STPMSI-TSISVFSTNMVIYFLFKFIIGFSRSQTSVALVLISERVSTRWR-PRATMI 215
: : | : : : : | : | : : : : | : : : : | : : | : :

Qy 235 VCTFFAVGMPLLPFAPIRD--WRMLLIALVPGVLCVPLWNFIPEPSRWLLISQRFR 291
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Db 216 PTFLEVLFGMSLSGIAFLAQDSWRVLYLTVSPAVFYCIYLIFALESRWLMHMOCKD 275
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Qy 292 EADIITQAAKNNNTAVPAVIFDSVEELNPKQQ-----KAFITDLDFRTNIAM 341
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Db 276 EADVLTKMSPKEK-----AYLESVYSKLPLEOENFEQAPTYSIKDFFRKWAFRLVV 330
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Qy 342 TMSLLLWMLTSGVFLSDAPNLHGDAYLNCFLSALIEPAYITAWLLLRTPRYII 401
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Db 331 MIIMFGL----GISYVGPLAARDIDVNYLSSETLNALVELPTFVTIPILLERNRSSV 386
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Qy 402 AAVLEFWGG--GVLLFIQLVPVDYFLSIGVLMGLKGFI TSA-----FSMLYVF 447
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Db 387 LVNTLGGASGVLCFV-----LSILGTEIAFAFELGTFECARIGFMLMAVF 433
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Qy 448 TAELPTLVNRWANGVISTASRVGSIATPYFVLGYAINRMPLPYVMGSLVLGIETLTF 507
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Db 434 MVSEMFPTCYRRSATMMERQALVWGGACCPLIASIGRIYPSVFAIFGAMSGLGMFVYL 493
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Qy 508 PESLGMTLPETELOMQK 524
: : | : | : : : : | : | : : : : | : : : : | : : | : :

Db 494 PETRGSLCDSNEEQEK 510
: : | : | : : : : | : | : : : : | : : : : | : : | : :

RESULT 12
B96825
hypothetical protein T8K14.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2000
C:Accession: B96825
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White-
ansen, N.F.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khayki,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maity,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun-
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719
A:Accession: B96825
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-515 <STO>
A:Cross-references: GB:AE005173; NID:g4835768; PIDN:AD30235.1; GSPDB:GN000
C:Genetics:
A:Gene: T8K14.17
A:Map position: 1

Query Match	14.5%	Score 412.5;	DB 2;	Length 521;
Best Local Similarity	28.4%	Pred. No. 4.9e-23;		
Matches 124;	Conservative	81;	Mismatches 179;	Indels 53; Gaps 12;

```

QY      117 WEFSQDVLSTVVTENLVCEDNKKVPLTTSLEFFVGVLGSFVGQL-SDRGRKN-VLF 174
      ||:  :||:| | :  :||:| | :  :||:| | :  :||:| | :  :||:| | :

```

RESULT 15

C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: F75580
 R:White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: F75580
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-454 <WHI>
 A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PID:AAF12486.1; PID:g646078
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0271
 A:Map position: 2
 C:Superfamily: yaaU protein

Query Match	11.5%	Score 327;	DB 2;	Length 454;
Best Local Similarity	26.0%	Pred. No. 8.8e-17;		
Matches 107;	Conservative 81;	Mismatches 166;	Indels 58;	Gaps 13;

QY 143 PLTTSLF---FVGLGSFVSGQLSDRGRKNVLFATMAYOTGFSFLOIF--SISHEM-- 195
 DB 67 PAATMLLTATFAGMLFGAWFGYLDADVGRRSVFLTTVALGVVFGLAGALAPTTLTLLVA 126
 QY 196 -FTVLFIIVMGQISNYVAFILGILGKSVRIIFSTLGVCTFFFAVGMMLPLFAYFI- 253
 DB 127 RLFTGFAIGGTLVDYSNMAEFVPTAMRGFLVYLES-----FWAGTVVVAALAWWVS 180
 QY 254 -----RDWRMLLALTPGVLCVPLWTFIPESPRWLSORRPREAEIIQKAAKNNTA 307
 DB 181 TAFAPAGRWLLGLAALPGVLGLIARIGIPDSRSLARGEAAQARAALQKVAQANGGT 240
 QY 308 VPAVIFDSVEELNPLKQKAPILDLFRTRIAIMTMSLLMLTSGVGFALSILDAPNL- 366
 DB 241 LPNAPL-----AHPDPPRVSPAQLFR-GVLARRTPLLMVTWFGLSLGYGIFSWLPSEL 294
 QY 367 -----HGDVYLNCPLSALIEIPAYITAWLLRLPRRYIIAAVLFWGG-GVLLFIQLVP 419
 DB 295 RAQGLDLGAVYRSTLLALAQVPGYLLAAYLVEKIGRRVTLVGFLTLGAVGAYLFLAHD 354
 QY 420 VDYYFLSIGVLMGKFGITSAFSLYVFTAEYPTLVNRNMAVGVTSTASRVGSIAPYFV 479
 DB 355 ANTVLLTSALL---\$FALLGAWGSLYAYTPELPTPLRTTGMGLVSGVARLASVWSP--- 408
 QY 480 YLGAYNRMLPYIVMGSLTVLIGIFTLFFP-----ESLGMTLPETLE 520
 DB 409 SIGA-----MLLTGNTLTALTFAVCFALAAALAAWGIQVETRQALAEAE 454

Search completed: January 25, 2002, 10:09:47
 Job time: 83 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:09:54 ; Search time 18.96 Seconds
(without alignments)
1065.523 Million cell updates/sec

Title: US-09-521-195-1
Perfect score: 2845
Sequence: 1 MRDYDEVIAFLGFWGPFQRL.....KKTRDSMETEENPKVLITAF 551

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2218	78.0	557	1 OCN2_HUMAN	O76082 homo sapien
2	2165	76.1	557	1 OCN2_RAT	O70594 rattus norv
3	2164	76.1	557	1 OCN2_MOUSE	Q92068 mus musculu
4	393	13.8	751	1 YLX5_CAEEL	P46501 caenorhabdi
5	317.5	11.2	1222	1 YMP3_CAEEL	Q10947 caenorhabdi
6	307	10.8	435	1 YR13_CAEEL	Q10917 caenorhabdi
7	294.5	10.4	529	1 YQ01_CAEEL	P30638 caenorhabdi
8	290	10.2	400	1 YCE1_BACSU	C34691 bacillus su
9	274.5	9.6	443	1 YRAU_ECOLI	P31679 escherichia
10	270	9.5	461	1 CSRC_BACSU	P46333 bacillus su
11	262	9.2	445	1 YGCS_ECOLI	Q46909 escherichia
12	248	8.7	490	1 GTR1_CHICK	P46896 gallus gall
13	247	8.7	592	1 HXT5_YEAST	P38695 saccharomyc
14	245	8.6	459	1 YDJK_ECOLI	P76230 escherichia
15	245	8.6	566	1 KHT2_KLUILA	P53387 kluyveromyc
16	244.5	8.6	567	1 HXT9_YEAST	P40885 saccharomyc
17	243	8.5	451	1 GTR1_PIG	P20303 sus scrofa
18	243	8.5	492	1 GTR1_RAT	P11167 rattus norv
19	240	8.4	495	1 GTR3_CANFA	P47842 canis famli
20	238	8.4	492	1 GTR1_BOVIN	P27674 bos taurus
21	238	8.4	492	1 GTR1_MOUSE	P17809 mus musculu
22	237.5	8.3	452	1 YDJE_ECOLI	P38055 escherichia
23	237	8.3	492	1 GTR1_HUMAN	P11166 homo sapien
24	237	8.3	546	1 HXT0_YEAST	P43581 saccharomyc
25	237	8.3	567	1 HXTA_YEAST	P54862 saccharomyc
26	236	8.3	451	1 YRAJ_BACSU	P37514 bacillus su
27	234.5	8.2	494	1 GTR3_SHEEP	P47843 ovis aries
28	234	8.2	491	1 XYLE_ECOLI	P09098 escherichia
29	231	8.1	522	1 GTR2_RAT	P12336 rattus norv
30	226	7.9	457	1 YIRO_YEAST	P40441 saccharomyc
31	225	7.9	492	1 GTR1_RABIT	P13355 oryctolagus
32	225	7.9	763	1 RGT2_YEAST	Q12300 saccharomyc
33	223	7.8	413	1 MUCK_ACICA	P94131 acinetobact

RESULT 1
OCN2_HUMAN
ID OCN2_HUMAN STANDARD; PRT; 557 AA.
AC O76082;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ORGANIC CATION/CARNITINE TRANSPORTER 2 (SOLUTE CARRIER FAMILY 22,
DE MEMBER 5) (HIGH-AFFINITY SODIUM-DEPENDENT CARNITINE COTRANSPORTER).
GN SLC22A5 OR OCTN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96289574; PubMed=9618255;
RA Wu X., Prasad P.D., Leibach F.H., Ganapathy V.;
RT "cDNA sequence, transport function, and genomic organization of human
RT OCTN2, a new member of the organic cation transporter family.";
RL Biochem. Biophys. Res. Commun. 246:589-595(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=9632077; PubMed=9685390;
RA Tamai I., Ohashi R., Nezu J.-I., Yabuuchi H., Oku A., Shimane M.,
RA Sai Y., Tsuji A.;
RT "Molecular and functional identification of sodium ion-dependent, high
RT affinity human carnitine transporter OCTN2.";
RL J. Biol. Chem. 273:20378-20382(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99113835; PubMed=9916797;
RA Nezu J., Tamai I., Oku A., Ohashi R., Yabuuchi H., Hashimoto N.,
RA Nikaide H., Sai Y., Koizumi A., Shoji Y., Takada G., Matsushita T.,
RA Yashino M., Kato H., Ohura T., Tsujimoto G., Hayakawa J., Shimane M.,
RA Tsuji A.;
RT "Primary systemic carnitine deficiency is caused by mutations in a
RT gene encoding sodium ion-dependent carnitine transporter.";
RL Nat. Genet. 21:91-94(1999).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=99384224; PubMed=10454528;
RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
RA Chen J., Conway S.J., Ganapathy V.;
RT "Functional characteristics and tissue distribution pattern of organic
RT cation transporter 2 (OCTN2), an organic cation/carnitine
RT transporter.";
RL J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
RN [5]
RP VARIANT CDSP GLN-169.
RX MEDLINE=99355597; PubMed=10425211;
RA Burwinkel B., Kreuder J., Schweitzer S., Vorgerd M., Gempel K.,
RA Gerbitz K.-D., Killmann M.W.;
RT "Carnitine transporter OCTN2 mutations in systemic primary carnitine
RT deficiency: a novel Arg169Gln mutation and a recurrent Arg282ter

ALIGNMENTS

mutation associated with an unconventional splicing abnormality.";
Biochem. Biophys. Res. Commun. 261:484-487(1999).
[6]
RN VARIANT CDSB CYS-211.
RP MEDLINE=99408248; PubMed=10480371;
RX Vaz F.M., Scholte H.R., Ruiters J., Hussaarts-Odijk L.M.,
RA Rodriques Pereira R., Schweitzer S., de Klerk J.B.C., Waterham H.R.,
RA Wanders R.J.A.;
RT "Identification of two novel mutations in OCTN2 of three patients with
RN systemic carnitine deficiency";
RN Hum. Genet. 105:157-161(1999).
[7]
RN VARIANT CDSB LEU-478.
RP MEDLINE=99172075; PubMed=10072434;
RX Tang N.L., Ganapathy V., Wu X., Seth P., Yuen P.M.,
RA Wanders R.J., Fok T.F., Hjelm M.H.;
RT "Mutations of OCTN2, an organic cation/carnitine transporter, lead to
RN deficient cellular carnitine uptake in primary carnitine deficiency";
RN Hum. Mol. Genet. 8:655-660(1999).
[8]
RN CHARACTERIZATION OF VARIANT CDSB LEU-478, AND MUTAGENESIS.
RP MEDLINE=20026865; PubMed=10559218;
RX Seth P., Wu X., Huang W., Leibach F.H., Ganapathy V.;
RT "Mutations in novel organic cation transporter (OCTN2), an organic
RN cation/carnitine transporter, with differential effects on the
RN organic cation transport function and the carnitine transport
RN function";
RN J. Biol. Chem. 274:33388-33392(1999).
[9]
RN VARIANTS CDSB ARG-283 AND PHE-446.
RP MEDLINE=20081068; PubMed=10612840;
RX Mayatepek E., Nezu J., Tamai I., Oku A., Katsura M., Shimane M.,
RA Tsuji A.;
RT "Two novel missense mutations of the OCTN2 gene (W283R and V446F) in a
RN patient with primary systemic carnitine deficiency";
RN Hum. Mutat. 15:118-118(2000).
[10]
RN VARIANT CDSB LYS-452.
RP MEDLINE=20145665; PubMed=10679939;
RX Wang Y., Kelly M.A., Cowan T.M., Longo N.;
RT "A missense mutation in the OCTN2 gene associated with residual
RN carnitine transport activity";
RN Hum. Mutat. 15:238-245(2000).
CC -1- FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE
CC TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
CC INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
CC CARNITINE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN KIDNEY, SKELETAL MUSCLE,
CC HEART AND PLACENTA.
CC -1- DISEASE: DEFECTS IN SLC22A5 ARE THE CAUSE OF SYSTEMIC PRIMARY
CC CARNITINE DEFICIENCY (CDSB). CDSB IS AN AUTOSOMAL RECESSIVE
CC DISORDER OF FATTY ACID OXIDATION CAUSED BY DEFECTIVE CARNITINE
CC TRANSPORT. PRESENT EARLY IN LIFE WITH HYPOKETOTIC HYPOGLYCEMIA AND
CC ACUTE METABOLIC DECOMPENSATION, OR LATER IN LIFE WITH SKELETAL
CC MYOPATHY OR CARDIOMYOPATHY.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
CC CATION SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF057164; AAC24828.1; -
DR EMBL; AB015050; BAA29023.1; -
DR EMBL; AB016625; BAA36712.1; -
DR MIM; 603377; -
DR MIM; 212140; -
DR InterPro; IPR003662; sub_trnsportr

DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE ORGANIC CATION/CARNITINE TRANSPORTER 2 (SOLUTE CARRIER FAMILY 22, MEMBER 5) (HIGH-AFFINITY SODIUM-DEPENDENT CARNITINE COTRANSPORTER).
 GN SLC22A5 OR OCTN2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Kidney;
 RX MEDLINE=99113835; PubMed=9916797;
 RA Nezu J., Tamai I., Oku A., Ohashi R., Yabuuchi H., Hashimoto N.,
 RA Nikaido H., Sai Y., Koizumi A., Shoji Y., Takada G., Matsushita T.,
 RA Yashino M., Kato H., Ohura T., Tsujimoto G., Hayakawa J., Shimane M.,
 RA Tsuji A.;
 RT "Primary systemic carnitine deficiency is caused by mutations in a
 RT gene encoding sodium ion-dependent carnitine transporter.";
 RL Nat. Genet. 21:91-94(1999).
 RN [2]
 RN SEQUENCE FROM N.A., AND VARIANT JVS ARG-352.
 RC STRAIN=C3H;
 RX MEDLINE=99057546; PubMed=9837751;
 RA Lu K., Nishimori H., Nakamura Y., Shima K., Kuwajima M.;
 RA "A missense mutation of mouse OCTN2, a sodium-dependent carnitine
 RT cotransporter, in the juvenile visceral steatosis mouse.";
 RL Biochem. Biophys. Res. Commun. 252:590-594(1998).
 RN [3]
 RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=99384224; PubMed=10454528;
 RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
 RA Chen J., Conway S.J., Ganapathy V.;
 RT "Functional characteristics and tissue distribution pattern of organic
 RT cation transporter 2 (OCTN2), an organic cation/carnitine
 RT transporter.";
 RL J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
 CC -I- FUNCTION: SODIUM-ION DEPENDENT, HIGH-AFFINITY CARNITINE
 CC TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
 CC INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
 CC CARNITINE.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -I- DISEASE: DEFECTS IN SLC22A5 ARE THE CAUSE OF JUVENILE VISCERAL
 CC STEATOSIS (JVS).
 CC -I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
 CC CATION SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB015800; BAA36590.1; -;
 CC EMBL; AF111425; AAC99787.1; -;
 CC EMBL; AF110417; AFD54060.1; -;
 CC MGI; MGI:1329012; SLC22a5.
 CC InterPro; IPR003662; sub.transprtr.
 CC Pfam; PF00083; sugar.tr.1.
 CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 KW Transport; Transmembrane; Glycoprotein; Disease mutation.
 FT TRANSMEM 21 41
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 198 218 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 258 278 POTENTIAL.
 FT TRANSMEM 342 362 POTENTIAL.
 FT TRANSMEM 374 394 POTENTIAL.
 FT TRANSMEM 407 427 POTENTIAL.
 FT TRANSMEM 431 451 POTENTIAL.

FT TRANSMEM 489 509 POTENTIAL.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 352 352 L -> R (IN JVS).
 SQ SEQUENCE 557 AA; 62779 MW; 6093F0EB9612B204 CRC64;
 Query Match 76.1%; Score 2164; DB 1; Length 557;
 Best Local Similarity 72.9%; Pred. No. 5.7e-134;
 Matches 406; Conservative 70; Mismatches 175; Indels 6; Gaps 3;
 QY 1 MDYDEVTAFLGEMGPORLIFFLLSASIIIPNGFNGMSVFLAGTPEHRCRYPDAANLSS 60
 DB 1 MDYDEVTAFLGEMGPORLIFFLLSASIIIPNGFNGMSVFLAGTPEHRCRYPDAANLSS 60
 QY 61 AWRNNSVPLRLRDGREGVPHSCSRYLATIANFSALEPGRDVLQDLQESCLDGEWFS 120
 DB 61 AWRNHSIPLETGDRQVQKCRRYRLATIANFSELGLEPGRDVLQDLQESCLDGEWYD 120
 QY 121 QDVYLTSTVVTENLVCDNWKVPLTTSLEFVGVLLGSFVSQSLSDRFGKKNVLFATMAVQ 180
 DB 121 KDVFSTIVTENDLVCKDDWKAPLTTSLEFVGVLLGMSFISQSLSDRFGKKNVLFATMAVQ 180
 QY 181 TGFSELFQIFSISWEMFTLVFVVGQISNVYVAFILGTTELKSKSVRIIFSTLGVCTFFA 240
 DB 181 TGFSELFQVFSVNFEMFTLVFVVGQISNVYVAFVLTGTEILSKSRIIFATLGVCIFFA 240
 QY 241 VGYMLLPFAFIRDRMLLALTPGVLCVPLVWETPESPRWLISQRRFRAEDTIQKA 300
 DB 241 FGFVPLPFAFIRDRMLLALTPGVLCALWFFIPESPRWLISQRRFRAEDTIQKA 300
 QY 301 AKMNNTPAVIFD--SYEELNPLKQKAFILDIFRTNIAIMTMSLLMLTSGVGYFA 358
 DB 301 AKINGIVAPSTIFDPSLQDLNSTKPOLHHYDLIRTNIRVITIMSIITLWLTISGVYFG 360
 QY 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLLTLPRIYITAAVLFWGGVGLFIOLV 418
 DB 361 LSLDTPNLHGDIYVNCFLAAVEPAVYVLAQLLQYLPRIYISAAFLGGSVLLFMQLV 420
 QY 419 PVDYVFLSIGLVMLGKFGITSAFSLMYFTAEYPTLVLRNMAVGTSTASRVGSIAPYF 478
 DB 421 PSELYFLTALVMVGKFGITSAFSLMYFTAEYPTLVLRNMAVGTSTASRVGSIAPYF 480
 QY 479 VYLGAYNMLPYVVGSLTVLIGITLFFEPESLGMTLPETLEOMOKVWFRSGK---KTR 535
 DB 481 VYLGAYNMLPYVVGSLTVLIGITLFFEPESLGMTLPETLEOMOKVWFRSGK---KTR 540
 QY 536 DSMETEENPKVL-ITAF 551
 DB 541 MQKDGEESPTVLKSTAF 557
 RESULT 4
 YLA5_CAEEL STANDARD; PRT; 751 AA.
 AC P46501;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHETICAL 84.8 KDA PROTEIN F23F12.5 IN CHROMOSOME III.
 GN F23F12.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Du 2.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; U12965; AAA20607.1; -.
DR WormPep; F23F12.5; CE01252.
DR InterPro; IPR002184; Srib.
DR InterPro; IPR003662; sub_trnsport.
DR Pfam; PF02175; Srib; 1.
DR Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 57 77 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 137 157 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
FT TRANSMEM 379 399 POTENTIAL.
FT TRANSMEM 410 430 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 515 535 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 583 603 POTENTIAL.
FT TRANSMEM 614 634 POTENTIAL.
FT TRANSMEM 678 698 POTENTIAL.
FT TRANSMEM 751 771 POTENTIAL.
SEQUENCE 751 AA; 84832 MW; A6C4F43540295EFC CRC64;

Query Match 13.8%; Score 393; DB 1; Length 751;
Best Local Similarity 27.0%; Pred. No. 1.5e-18;
Matches 119; Conservative 91; Mismatches 184; Indels 46; Gaps 14;

QY 127 TVVTEWNLVCEDN-WKVPILTSFFVGVLLGSFVSQSLDRFGKKNVLPATMAVQGFSEF 185
DB 303 SMVODFMFCGRKAYDAAWATIQFVGLGATYHGLDHFGRKPKVPFSGISVILFV 362
QY 186 LQFISISWEMFTLVFVIGMGQISNVVAFILGTILGKSVRIIFSLGTGVCFFAVGY-- 243
DB 363 ASGPAPSEWFAFRFVIGHSIASILIVFAYILEFIEPEQVFLR-----SFFNNGYAR 417
QY 244 MLLPLFAYFIRDRMILLALTVPGLCVPLWFPFIPESPRWLISORRFREAEIIQKAAKM 303
DB 418 LVFTLACFICGYWRAAIAATSLSLPLVLLIPESPKWNTKKRPDARAAAEKRVWL 477
QY 304 NNTAVPAV-----IFDSVEELNPLKQOKAFIL-DLFRTRNIAITMISLLMLTSV 354
DB 478 --SGIPYNDQDSIEISEKLEE---KSTKIYTMKDLFTSWTIAYRTIVVGSLSFSTSL 531
QY 355 GYFALSIDAPNLGHDAVNCFLSALIEIPAYITAWLLRLTP---RRIYI-----IAAVLF 406
DB 532 SAFGSDNLGNLGNFVLSQFVSAVTAFAKIFVFLDLTVPSDRRLHQYQIAMILC 591
QY 407 WGGVGLLFQILVPV-----DYFLSLGLVLMGKFGITSAFSMLYVFTAEIPLVLR 457
DB 592 Y--CVIMVLMILPESDCSGSQRDLAIITINIGVSFEIT--WDACYLVAVECFPKIR 647
QY 458 NNAVGVTSASRGVSIATPYFVYLGANRMLPYIVMGSL-TVLGIGITFLFPESLGMTLP 516
DB 648 TIGIGTCLLARTGALLAPQIMAYISDIYRPAPYAVVCSIGTISLISCVLPDPKGVDL- 706
QY 517 ETLEQMKVWFRGSKTRD 536
DB 707 AALDPTEELDYRKKSMTEN 726

RESULT 5

YMP3_CAEEL
ID YMP3_CAEEL STANDARD; PRT; 1222 AA.
AC Q10947; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 139.9 KDA PROTEIN B0361.3 IN CHROMOSOME III.
GN B0361.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Du Z.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

EMBL; U00031; AAK18864.1; -.
DR WormPep; B0361.3; CE00752.
DR Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 232 252 POTENTIAL.
FT TRANSMEM 263 283 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 355 375 POTENTIAL.
FT TRANSMEM 382 402 POTENTIAL.
FT TRANSMEM 413 433 POTENTIAL.
FT TRANSMEM 442 462 POTENTIAL.
FT TRANSMEM 1222 AA; 139868 MW; CBA42A80A254FB80 CRC64;
SEQUENCE 1222 AA; 139868 MW; CBA42A80A254FB80 CRC64;

Query Match 11.2%; Score 317.5; DB 1; Length 1222;
Best Local Similarity 20.3%; Pred. No. 2e-13;
Matches 114; Conservative 106; Mismatches 219; Indels 123; Gaps 14;

QY 3 DYDEVIAFLGEGWGP-----QRLIFPLLSASIPNGFNGMSVVFAGTPEHRCRVPD 54
DB 35 DPKDFEAYGAYGKYQIFTYVLVQTLNFFYSSSMVI-----MSFVQL--NLEKQCE--- 83
QY 55 AANLSSAWNNNSVPLRLRGREPHSCSYRLATIANFALGLEPGRVDVLQLEQESCL 114
DB 84 -----YKNETIP-----ISETCQ-----IETESSKAFNLNGEYC- 113
QY 115 DGWEFSQDYL-----STVVTEWNLVCEDNWKVPLTTSFFVGVLLGSFVSQSLDR 166
DB 114 ---GIANTLVNVTNOKASTNLLVDFDLSCSHWFQEFGLTIFTIGAVTAVFMSMLADR 170
QY 167 FGKKNYLFATMAVQGFSEFLOIFSIISWEMFTLVFVIGMGQISNVVAFILGTILGKSV 226
DB 171 YGRKPIIVTTAILAFLANNAAGSPNFAIFLILRAFIGACSDSYLSVASVATCEYLSEKA 230
QY 227 RIIFSLGVCITFAVGYMLLPFAFIRDRMILLALTVPGLCVPLWFPFIPESPRWLIS 286
DB 231 R-AWITVYVYVAVMSLGMVWTLVLTMTDDWRNYFIVSLPGVYGFMALTFPESPHMLIT 289
QY 287 ORRFEAEADIIOKAAKMNNTAVPAVIFDSVEELNPLKQOKAFILDLFRTRNIAITMISL 346
DB 290 KNTEKLKVIKTANRM-----VISL 310

QY 347 LLMMLTSGVGFALSDAPNLHGDAYLNCFLSALIEIPAYITAMLLRLTLPRRYIIAAVL 406
Db 311 V-----YFAISFMSVGLGGDOVQAFLYSSLLIEIPAGLAVIPLMKWGRKMIVTWCLV 362
QY 407 WGG-----GVLLFTQLQVDPVYFISGLVLMGKFGITSAFSMLYVFTAEIYPTLVRNMAVG 462
Db 363 FOTLALIGVTVFL-----DSYEFKVLMLVAKVMATIIYSVHPWATEQFPITSVRSCLFS 417
QY 463 VTSTASVSGSIAPFYVYLGAYNRMPLPYIVMGSLTVLIGIFTLFFPESLGMILPETLEOM 522
Db 418 LMIIPQSGMIINGPYKHIYVFNWVIALFSPISATLAPMLHETKNKLPDIETLSL 477
QY 523 -----QKVMFRSGKKTDSM 538
Db 478 SYPSETNDLSAYRRSKSSSSV 499

RESULT 6

YU13_CAEEL STANDARD; PRT; 435 AA.
AC Q10917;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 48.6 KDA PROTEIN B0252.3 IN CHROMOSOME II.
GN B0252.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Du 2., Waterston R.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U23453; AAC46757.1;
DR WormPep; B0252.3; CE02419.
DR InterPro; IPR003662; sub_trnsportr.
DR Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 55 75
FT TRANSMEM 384 404
FT TRANSMEM 435 AA; 48571 MW; 5F6160359FA1840B CRC64;
SQ SEQUENCE 435 AA; 10.8%; Score 307; DB 1; Length 435;
Query Match
Best Local Similarity 23.5%; Pred. No. 3.3e-13;
Matches 94; Conservative 73; Mismatches 155; Indels 78; Gaps 8;

QY 128 VVTENLVNCDNWKVPLTSLFFVGLGVSGVSGQLSDRGKKNVLPATMAVQTFGSFLQ 187
Db 91 VADEFDLTGDSWLAEITTFYVWGNMIGMFIPPLADHYGLRPLVFNATVLLMVGGMIS 150
QY 188 IFSISWEMFTVLVIVMGQISVNYVAEILTEILGKSVRIIFSTLGVCTFFAVGYMLLP 247
Db 151 AFSTIMMFCIMRMTHGIFYTAAGLAGVLYGENTPLRLR-FTTSVYGVNVMVYGACFLG 209
QY 248 LFAYFIWRMLLALTPGV-LCVPLWWEIPESPRWLISORREREADIQK-AAKWN 305
Db 210 LLAYILPWRILMFCISVNPVIALIYMTVPESLHVLSSQQNEKTEAWLEKIRGPKGD 269
QY 306 TAVPAVIFDSVEELNPLK-----QOKAFILDIFRTRNIAIMTISLLMLTSLVGYFALS 361

Db 270 ISASDIVEDRNGSSFKTLCKREIKFTKTLFQ----- 302
QY 362 DAPNLHGDAYLNCFLSALIEIPAYITAMLLRLTLPRRYIIAAVLFWGGVLLFIQLVPVD 421
Db 303 -----DRIYI-----YLFWIGILLYEGRKPL- 324
QY 422 YFISIGLVLMGKFGITSAFSMLYVFTAEIYPTLVRNMAVGVTASVSGSIAPFYVYL 481
Db 325 FEFLA-----HEFGRSSSLHFFSDFHEQIFPTDGRNKCIGFCETLSRFGMLSPYLSHL 378
QY 482 GAYNRMPLPYIVMGSLTVLIGIFTLFFPESLGMILPETLEQ 521
Db 379 TAVHALAPAYITLSLIANGVGLLTLILPETLNTKLPSTIAE 418

RESULT 7

YU01_CAEEL STANDARD; PRT; 529 AA.
AC P30638; O21101;
DT 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 58.3 KDA PROTEIN ZK637.1 IN CHROMOSOME III.
GN ZK637.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MEDLINE=92168156; PubMed=1538779;
RA Sulston J., Du Z., Thomas K., Wilson R., Hillier L., Staden R.,
RA Halloran N., Green P., Thierly-Mieg J., Qiu L., Dear S., Coulson A.,
RA Craxton M., Durbin R.K., Berks M., Metzstein M., Hawkins T.,
RA Ainscough R., Waterston R.;
RT "The C. elegans genome sequencing project: a beginning."
RL Nature 356:37-41(1992).
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Durbin R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

CC EMBL; Z11115; CAA77460.1;
DR EMBL; Z22175; CAA77460.1; JOINED.
DR EMBL; Z22175; CAA80131.1;
DR EMBL; Z11115; CAA80131.1; JOINED.
DR PIR; S15786; S15786.
DR WormPep; ZK637.1; CE06638.
DR InterPro; IPR003662; sub_trnsportr.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Hypothetical protein; Transmembrane; Transport.
FT TRANSMEM 86 106
FT TRANSMEM 122 142
FT TRANSMEM 158 178
FT TRANSMEM 238 258
FT TRANSMEM 320 340
FT TRANSMEM 373 393
FT TRANSMEM 411 431
FT TRANSMEM 482 502
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.

Db	393	VLAQGMQVILLALLVLLGVVTVLWAPETKALPL	429	
RESULT	12			
GTRL_CHICK		STANDARD;	PRT;	490 AA.
AC	P46896;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	GLUCOSE TRANSPORTER TYPE 1 (GTL1).			
GN	SUC2AI OR GLUT1.			
OS	Gallus gallus (Chicken).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianinae;			
OC	Gallus.			
OX	NCBI_TaxID=9031;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96157892; PubMed=8589457;			
RA	Wagstaff P., Kang H.Y., Mylott D., Robbins P.J., White M.K.;			
RT	"Characterization of the avian GLUT1 glucose transporter:			
RT	differential regulation of GLUT1 and GLUT3 in chicken embryo			
RT	fibroblasts."			
RL	Mol. Biol. Cell 6:1575-1589 (1995).			
CC	- - FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM MAY BE			
CC	RESPONSIBLE FOR CONSTITUTIVE OR BASAL GLUCOSE UPTAKE. HAS A VERY			
CC	BROAD SUBSTRATE SPECIFICITY; CAN TRANSPORT A WIDE RANGE OF ALDOSES			
CC	INCLUDING BOTH PENTOSE AND HEXOSE (BY SIMILARITY).			
CC	- - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	- - SIMILARITY: HIGH TO OTHER GLUCOSE TRANSPORTERS.			
CC	- - SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
DR	EMBL; L07300; AAB02037.1;			
DR	InterPro; IPR003663; Sugar transportr.			
DR	InterPro; IPR003662; sub_transportr.			
DR	Pfam; PF00083; sugar_tr; 1.			
DR	PRINTS; PRO0171; SUGTRNSPORT.			
DR	PRINTS; PRO0172; GLUCTRNSPORT.			
DR	PRINTS; PRO1190; GLUCTRSPORT1.			
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
KW	Transmembrane; Sugar transport; Transport; Glycoprotein;			
KW	Multigene family.			
FT	TRANSMEM 12	32	POTENTIAL.	
FT	DOMAIN 33	55	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM 66	86	POTENTIAL.	
FT	TRANSMEM 95	115	POTENTIAL.	
FT	TRANSMEM 126	146	POTENTIAL.	
FT	TRANSMEM 155	175	POTENTIAL.	
FT	TRANSMEM 185	205	POTENTIAL.	
FT	DOMAIN 206	270	CYTOPASMIC (POTENTIAL).	
FT	TRANSMEM 271	291	POTENTIAL.	
FT	TRANSMEM 306	326	POTENTIAL.	
FT	TRANSMEM 337	357	POTENTIAL.	
FT	TRANSMEM 368	388	POTENTIAL.	
FT	TRANSMEM 401	421	POTENTIAL.	
FT	TRANSMEM 429	449	POTENTIAL.	
FT	DOMAIN 450	490	CYTOPASMIC (POTENTIAL).	
FT	CARBOHYD 44	44	N-LINKED (GLCNAC...) (POTENTIAL).	
SQ	SEQUENCE 490 AA;	54086 MW;	63BB176812D06104 CRC64;	
Query Match		8.7%;	Score 248;	DB 1; Length 490;

Best Local Similarity	24.2%;	Pred. No. 2.6e-09;				
Matches 117;	Conservative 87;	Mismatches 170;	Indels 110;	Gaps 21;		
QY	117	WFEFSDVYLSTVTWNLVCEDNWKPVLTTSLFFVGVLLGSFVSQSLDRGRK	-----	170		
Db	49	YRYEERIPATLTLWS	-----LSVAIFSVGGMIGSFVGLFVNRGRNRMSMLS	98		
QY	171	NVLPMATMAVOTGFSFLQIFSISWEM	-----FTVLFVIVGMGOISNYVAFIL	217		
Db	99	NILAFALAAVLMGFSKM	-----ALSFEMLILGRFLIGLYSLTTGTFVPMYGEVSPALRGAL	155		
QY	218	GTEILGKSVRIIFSTLGVCTFFAVGYMLLPFAFYI	---RD---WRMLLLALTVPGLVCP	272		
Db	156	GT	-----FHOLGI	-----VLGILTAQVFGDLNMGDSLMLPLLLGFIIVFALLQCI	201	
QY	273	LWMTFIPESPRW	---LISQRRFEAEDIQKAARNNTAVPAVIFDSVEELNP	---LKQOKAFIL	330	
Db	202	ILPFAPESPRELLINRNEENAKSVLKKL	---RGITDVSSDLQEMKEESQOMREKKVTIM	259		
QY	331	DLFRTRN	-----IAIMTMSLLMLTSVGYFALSADPNLHGDAYLNCFLSALIEIPA	384		
Db	260	ELFRSPMYRQPIILAIIVLQLSQOLSGINAVFYTSI	-----FEKSGVEQPV	306		
QY	385	YIT	-----AWLLRLTLPRR	---YIIAAVLFWGGVLLFIQIOLVPDYY	---FL	425
Db	307	YATIGSGVVNTAFTVVSILFVVERAGRRTLHLIGLAGMAGCAILMTIATLDDQPMWSYL	366			
QY	426	SIGLVMLKFGITSAFS	-----MLYFTAEIYPTLVNRMAVGTSTASRVGS	---IIAPYFV	479	
Db	367	SIVAI	---FGVAFIEIGPGPIPFWIVAEIQLFSGPRPAFAVAGLSNWTSTNFIVGMGFO	422		
QY	480	YLGAYNRMLPYIVMGSITVLIGITLFP	---FPESLGMTLPETLEQMKVVKFSGSKKTRDSM	538		
Db	423	YIAQLGSGVFIITVLVLVLIFFIYFKVPETKGTDEIAYR	-----FROGASQSDK	476		
QY	539	ETEE 542				
Db	477	TPDE 480				
RESULT	13					
HXT5_YEAST		STANDARD;	PRT;	592 AA.		
AC	P38695;					
DT	01-FEB-1995 (Rel. 31, Created)					
DT	01-FEB-1995 (Rel. 31, Last sequence update)					
DT	01-FEB-1995 (Rel. 31, Last annotation update)					
DE	PROBABLE GLUCOSE TRANSPORTER HXT5.					
GN	HXT5 OR YHR096C.					
OS	Saccharomyces cerevisiae (Baker's yeast).					
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;					
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.					
OX	NCBI_TaxID=4932;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=MC996;					
RA	Reifenberger E., Koetter P., Ciriacy M.;					
RL	Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=S288C / AB972;					
RX	MEDLINE=94378003; PubMed=8091229;					
RA	Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,					
RA	Du Z., Favella A., Fulton L., Gattung S., Geisel C., Kirsten J.,					
RA	Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,					
RA	Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,					
RA	Nhan M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,					
RA	Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,					
RA	Vaudin M.;					
RT	"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome					
RT	VIII."					
RL	Science 265:2077-2082(1994).					
CC	- - FUNCTION: PROBABLE GLUCOSE TRANSPORTER.					

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- MISCELLANEOUS: GLUCOSE TRANSPORT IS THOUGHT TO BE MEDIATED BY TWO
 CC KINETICALLY DISTINCT SYSTEMS, A GLUCOSE-REPRESSIBLE HIGH-AFFINITY
 CC SYSTEM AND A CONSTITUTIVE LOW-AFFINITY SYSTEM.
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X77961; CA54923.1; -;
 CC EMBL; U00060; AAB68934.1; -;
 CC PIR; S43742; S43742.
 CC PIR; S46726; S46726.
 CC SGD; S0001138; HXT5.
 CC InterPro; IPR003663; Sugar_trnsportr.
 CC InterPro; IPR003662; sub_trnsportr.
 CC Pfam; PF00083; sugar_tr; 1.
 CC PRINTS; PR00171; SUGTRNSPORT.
 CC PRINTS; PR00172; GLUCTRNSPORT.
 CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 CC Repeat; Transmembrane; Sugar transport; Glycoprotein.
 KW DOMAIN 1 82 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 83 103 POTENTIAL.
 FT DOMAIN 104 137 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 138 158 POTENTIAL.
 FT DOMAIN 159 164 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 165 185 POTENTIAL.
 FT DOMAIN 186 195 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 196 216 POTENTIAL.
 FT DOMAIN 217 222 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 223 243 POTENTIAL.
 FT DOMAIN 244 257 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 258 278 POTENTIAL.
 FT DOMAIN 279 380 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 381 401 POTENTIAL.
 FT DOMAIN 402 409 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 410 430 POTENTIAL.
 FT DOMAIN 431 450 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 451 471 POTENTIAL.
 FT DOMAIN 472 487 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 488 508 POTENTIAL.
 FT DOMAIN 509 512 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 513 533 POTENTIAL.
 FT DOMAIN 534 592 EXTRACELLULAR (POTENTIAL).
 FT SEQUENCE 592 AA; 66251 MW; 6906721BAC1A5F87 CRC64;
 Query Match 8.7%; Score 247; DB 1; Length 592;
 Best Local Similarity 22.9%; Pred. No. 3.6e-09;
 Matches 114; Conservative 70; Mismatches 172; Indels 142; Gaps 22;
 Qy 124 YLSTVTVENLVCDNKKVPLTSLFFVGVLLGSFVSQGLSDRGKRVLFATMAVQTGF 183
 Db 130 YLSDV-----RTGLMVSIFNGCAIGGIVLSKLGDMYGRKIGLMTVVVVISG 177
 Qy 184 SFLQIFST-SWEMFTLVFVVGMOISNYYVAFTLGPILGKSVRIIFSTLGVCTFFAVG 242
 Db 178 IIIQIASIDKKWQYFIRIGISGLGVGIGITVLAPMLISEVSPKQLR---GLVSC----- 228
 Qy 243 YMLLPLFAFYTR-----DWRMLLLALTVPGLVCLPLMW-----FIPE 279
 Db 229 YQLMTTFGIFLGYCTNFGTKNSNSVQWR-----VPLGLCFA-WSIFMIVGMFTVPE 279
 Qy 280 SPRLWISORRPREAEDIIQAKNNNTAVPAVFD-----SVE-----ELNPL 322
 Db 280 SPRLVLEVKGTEAKRSLRANKTTDS-PLVTLEMYNQSSIEAERLAGSASWGLVTVG 338

Qy 323 KQAKAFILDPRTRNIAITMTSLMLLWMLTSVGYF-----ALSDAPNLHGDAYLN 373
 Db 339 KPQ-----MFRF---TLGMMISSLQOOLGDNFFYTGTTFOAVGLE-----DSFET 383
 Qy 374 CFLSALIEIPAYITAWLLRLPRRYIAAFLFWG--GGVLLFIQLVPVDYFSLSIGLVM 431
 Db 384 AIVLGVVN--FVSTFSLTYVD-RFGRNCLLWGCVMICCV-----VVASVGVTR 432
 Qy 432 LKRGF-----ITSAFSMLYVFTAEIPLPLVRNMAVGVTSTA 467
 Db 433 LWPNGQDQSPSKGAGNCMIVFACFYIFCFATWAPVAVYLISESYPLVRKAMSIASAC 492
 Qy 468 SRVGSIIAPYFV--YLGAYNRMLPYIVMGSLLTVLIGITFTFFPESLGMTPTLEQMQK- 524
 Db 493 NWINGFLISFTPTITSAINFYGYGVFMCVMFAFYVFFVFPETKGLTLEVNMEYEN 552
 Qy 525 -----VKWFRSGKKTRD 536
 Db 553 VLPWKSTKWIPPSRRTD 570
 RESULT 14
 YDJK_ECOLI
 ID YDJK_ECOLI STANDARD; PRT; 459 AA.
 AC P76230; P76911;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YDJK.
 GN YDJK OR B1775.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sasaki G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 Genome
 RT Corresponding to the 28.0-40.1 min Region on the Linkage Map.";
 RL DNA Res. 3:363-377(1996).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000272; AAC74845.1; -;
 CC EMBL; D90821; BAA15573.1; ALT_INIT.
 CC EcoGene; EGI3487; ydjk.
 CC InterPro; IPR003662; sub_trnsportr.

```

DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
FT TRANSMEM 113 133 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 272 292 POTENTIAL.
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
SQ SEQUENCE 459 AA; 49602 MW; 89AB53E8BDD77 CRC64;

Query Match 8.6%; Score 245; DB 1; Length 459;
Best Local Similarity 23.4%; Pred No. 3.7e-09;
Matches 97; Conservative 81; Mismatches 177; Indels 60; Gaps 17;

QY 134 LVCEDN-----WKVPLTSLF-----FVGLGSGVSQLSDRGRKNVLF 175
DB 34 LVCNNAVAGGLILAQKALGWTDNSTATFSALTITAGMFGALVGLIGDKTGRNAFL 93
QY 176 TMAVQTGFSLQIFSI-SWEMFTVLFVIVGMQISNVVAFILGTEIL-GK-----SVRII 229
DB 94 YEAIHTASVVGAFSPNMFDFLIACRFVGVGLGALLVTLFAGFTFYMPPGRNRTWSSRVS 153
QY 230 FS---TLGVCTFAVGMMLPLFAFIRDRWMLLALTVPGLVCPPL-WWFIPESPRWLI 285
DB 154 FIGNWYPLCSLIANG--LTPLIS-AENWVRQLLIPALISLIATLALAWRPESPRWLE 210
QY 286 SQRFPREAEITQK-----AKNNNTAVPAVIFDSVEELNPLKQKAFILDLFRTRI--- 338
DB 211 SGRYQAEKVRMSIEEGVIRGTGKPLPPVVIADBGKAPQAVPYSAITLVGLLKRVLGS 270
QY 339 AINTIMSLLLMLTSGVFEALSIDAPNLHGDAYLNCFLSALIEIP-AYTAWLLLRTPR 397
DB 271 CVLIANNVQVYTLINWLPFIEMTQGLNKDSIVLNTM--SMGAPFGFIAMLVMDKIPR 328
QY 398 RYIIAAVFWGGVLLFIQLVPVDYFYS--TGLVMLGKFGITSAFSMLY-----VFTA 449
DB 329 KTM-----GVGLILIAVLGYIYSLOTSMLLTILGIFLIT--FVYMVYCYASAVVP 379
QY 450 ELYPTLVNRNAVGVSTASRVGSIIAPY--FVYLGAYNRMLPYIVGSLTVLIGI 502
DB 380 EIWPTKLRGSLGSLANVGRISGIAAPYANAVALLSSYGVTVGTVILLGAVSIIAVI 434

RESULT 15
KHT2_KLULA
ID KHT2_KLULA STANDARD; PRT; 566 AA.
AC P53387.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HEXOSE TRANSPORTER 2.
GN KHT2.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAG6.
RX MEDLINE=98028406; PubMed=9363776;
RA Weirich J., Goffinl P., Kuger P., Ferrero I., Breunig K.D.;
RT "Influence of mutations in hexose-transporter genes on glucose
repression in Kluyveromyces lactis."

```

```

RL Eur. J. Biochem. 249:248-257(1997).
CC -!- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z47080; CAA87389.1; -.
DR InterPro; IPR003663; Sugar_transprot.
DR InterPro; IPR003662; sub_transprot.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PRINTS; PR00172; GLUCTRNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Repeat; Transmembrane; Sugar transport; Transport; Glycoprotein.
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 169 189 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 234 254 POTENTIAL.
FT TRANSMEM 357 377 POTENTIAL.
FT TRANSMEM 386 406 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 465 485 POTENTIAL.
FT TRANSMEM 491 511 POTENTIAL.
SQ SEQUENCE 566 AA; 62727 MW; 9784173EC2375444 CRC64;

Query Match 8.6%; Score 245; DB 1; Length 566;
Best Local Similarity 22.2%; Pred. No. 4.7e-09;
Matches 106; Conservative 78; Mismatches 196; Indels 98; Gaps 17;

QY 139 NWKVPILTSLFFVGVVLGSGVSGQLSDRGRKNVLPATMAVQTGFSLQIFSI-SWEMFT 197
DB 109 NVRTGLIVSFNIGCAIGGIILSKLGDYGRRIGLMTWLVVYVVGIIQIASIDKWKVQYF 168
QY 198 VLFVIVGMQISNVVAFILGTEILGKSVRIIFSTLGVCTFAVGMMLPLFAVFI--- 254
DB 169 IGRISGLGVGGISVLSFMLISETAPKHIR-----GLVSP---YQLMITFGIFLYCT 219
QY 255 -----DWRMLLALTVPGLVCPPLMWFPFIPESPRWLI-SQRFPREAEITQKAKM 303
DB 220 NYGTFTKYSNSVQWRVPLGLCFAWAIFMTGMLFVPSPRFLVEKDRIDEAKRSIAKSNV 279
QY 304 N--NTAVPA--VIFDSVEELNPLKQKAFILDLFRTR-NIAIMTMSLLL---WMLTSV 354
DB 280 SYEDPAVQAEVDLICAGVEA--ERLAGSASIKELFSTKTKVFORLIMGMLIQSFOOLTGN 337
QY 355 GYF-----ALSLDAPNLHGDAYLNCFLSALIEIPAYITAWLLLRTPRPIIAAVL 405
DB 338 NYFPFYGTTFISVSGMD-----DSFETSVLGVINFASTFVAIYVVDKFGRR---KCL 387
QY 406 FWGGVLLFIQLVPVDYFYSIGLVMGLKFG-----ITS 439
DB 388 LWGAAMTACMVV-----FASVGVTRLPDGDANHPETASKGAGNCMIVFACFYIFCFATS 442
QY 440 AFSLMYVFTAEPLYTLVRNNAVGVSTASRVGSIIAPYFV--YLGAYNRMLPYIVGSLT 497
DB 443 WAPIAYVVVAESYPLRVKAKMAIATASNWIWGFINGFTPTTSIAHFYGYGVFMGCLV 502
QY 498 VLIGITFLFPFESLGMTLPTEQMOK-----VKWFRSGKKTR---DSMETEENP 544
DB 503 AMFFYVFFVFPETKGLTLEEVQEMWEEGVLPWKSSSWPSSRRNAGYDVALQHDKEP 560

```

Search completed: January 25, 2002, 10:15:19
Job time: 325 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:09:19 ; Search time 47.78 Seconds
(without alignments)
1686.816 Million cell updates/sec

Title: US-09-521-195-1

Perfect score: 2845
Sequence: 1 MRDYDEVIAFLGEGWPFQRL.....KKTRDSMETEENPKVLITAF 551

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2845	100.0	551	4	O14546
2	2833	99.6	551	4	Q9H015
3	2484	87.3	553	11	Q9R141
4	2470	86.8	553	11	Q9Z306
5	2034.5	71.5	564	11	Q9WTN6
6	857	30.1	548	5	Q9VCA2
7	852	29.9	548	5	O01384
8	819	28.8	567	5	Q9VCA3
9	773.5	27.2	561	5	Q9V6L6
10	754.5	26.5	568	5	Q9U539
11	754.5	26.5	576	5	O02270
12	707	24.9	554	6	O77504
13	703	24.7	555	4	O15244
14	700	24.6	554	6	O02713
15	699.5	24.6	593	11	P70485
16	696.5	24.5	553	11	O70577
17	695.5	24.4	555	11	Q9R0W2
18	695	24.4	674	5	Q9V1K2
19	692	24.3	556	11	O08966

20	691.5	24.3	593	11	P97558
21	688.5	24.2	554	4	O15395
22	688.5	24.2	554	4	Q9NOD4
23	687.5	24.2	556	11	Q63089
24	687	24.1	556	11	Q9R1Q4
25	684.5	24.1	554	4	O15245
26	673	23.7	535	11	Q63314
27	672	23.6	557	5	Q9VEX8
28	671	23.6	556	4	O75751
29	653.5	23.0	456	4	O14567
30	653.5	23.0	548	4	Q9Y694
31	652	22.9	539	4	Q9H2W5
32	650	22.8	551	11	Q9WTW5
33	648.5	22.8	575	5	Q9BKR9
34	648	22.8	551	4	O9Y226
35	648	22.8	551	11	Q99JF0
36	640	22.5	551	11	O88446
37	638	22.4	537	11	O88909
38	637	22.4	536	11	Q9R1U7
39	627	22.0	538	5	Q9V6H5
40	618.5	21.7	562	13	O57379
41	604.5	21.2	545	11	O61185
42	602.5	21.2	551	11	O35956
43	599	21.1	553	11	O54778
44	592	20.8	762	5	Q9VNW8
45	586	20.6	542	5	Q9VNX2

ALIGNMENTS

RESULT	1
O14546	
ID	O14546
AC	O14546; PRELIMINARY; PRT; 551 AA.
DT	01-JAN-1998 (Tremblrel. 05, Created)
DT	01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE	POLYSPECIFIC ORAGANIC CATION TRANSPORTER.
DE	OCN1.
GN	OCN1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RP	TISSUE=LIVER;
RX	MEDLINE=98086199; PubMed=9426230;
RA	Tamai I., Yabuuchi H., Nezu J., Sai Y., Oku A., Shimane M., Tsuji A.;
RT	"Cloning and characterization of a novel human pH-dependent organic cation transporter, OCTN1.";
RT	FEB Lett. 419:107-111(1997).
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC	-!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR	EMBL; AB007448; BAA23356.1; -
DR	InterPro; IPR003662; sub.transporter.
DR	Pfam; PF00083; sugar.tr.1
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW	Transmembrane.
SQ	SEQUENCE 551 AA; 62176 MW; F5903421C789F60A CRC64;

Query Match	100.0%; Score 2845; DB 4; Length 551;
Best Local Similarity	100.0%; Pred. No. 6.5e-167; Indels 0; Gaps 0;
Matches 551; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MRDYDEVIAFLGEGWPFQRLIFLLSASTIPNGFMGMSVFLAGTPEHRCRVPDAANLSS 60
Db	1 MRDYDEVIAFLGEGWPFQRLIFLLSASTIPNGFMGMSVFLAGTPEHRCRVPDAANLSS 60
QY	61 AWRNNSVPLRLRDGREVPHSCSRYLATIANFSAIGLEPGRDVDLGQLESCLDGWFEFS 120
Db	61 AWRNNSVPLRLRDGREVPHSCSRYLATIANFSAIGLEPGRDVDLGQLESCLDGWFEFS 120


```
QY 121 QDYVLSVTWTEWNLVCEENKWKPLTTSLSFFVGVLLGSFVSQGLSDRFGKKNVLFATMAVQ 180
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 KDVELSTIVTEWNLVCEENKWKPLTTSLSFFVGVLLGSFVSQGLSDRFGKKNVLFATMAVQ 180
QY 181 TGFSEFQIFSIWEMFTVLFVIVGMQISNYVAVFTLGTTEILGKSVRIIFSTLGVCTFFA 240
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 TGFSEFQIFSNWEMFTVLFVIVGMQISNYVAVFTLGTTEILGKSVRIIFSTLGVCTFFA 240
QY 241 VGYMLPLFAFYFIRDRWMLLLALTVPGLVPLWMTIPESPRWLSIORRPREAEDIQKA 300
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 IGYMVLPLFAFYFIRDRWMLLLALTVPGLVPLWMTIPESPRWLSIORRPREAEDIQKA 300
QY 301 AKMNTAVPAVIEDSVE--ELNPLKQOKAFILDLFTRNIAIMTMSLLWMLTSVGYFA 358
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 AKMNGIMAPVIEDPDELQELNSLKQKVFILDLFTRNIAITVMSVWMLWMLTSVGYFA 360
QY 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWMLLRLTPRRYIIAAVLFWGGVLLFIQIV 418
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 361 LSLNPNLHGDAYLNCFLSGLIEVPAYITAWMLLRLTPRRYIIAGVLFWGGVLLFIQIV 420
QY 419 PVDYFISIGLVMKFGITSAFMSLYVFTAEYPTLVNRNAVGVTSASRVGSIITAPYF 478
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 421 PEDYNFYSIGLVMKFGITSAFMSLYVFTAEYPTLVNRNAVGVTSASRVGSIITAPYF 480
QY 479 VYLGAYNRLPYIVMGLSVTLVIGLFTLFPESLGMTLPETLEOMQKVKWFRSGKKTDRSM 538
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 481 VYLGAYNRLPYIVMGLSVTLVIGLFTLFPESLGMTLPETLEOMQKVKWFRSGKKTDRSM 540
QY 539 ETEENPKVLITAF 551
:||||:||||:
Db 541 DREENPKVLITAF 553
```

RESULT 4

```
Q92306 Q92306 PRELIMINARY; PRT; 553 AA.
ID Q92306;
AC Q92306;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DE ORGANIC CATION TRANSPORTER.
GN SLC22A4 OR OCTN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL 6J; TISSUE=KIDNEY;
RA Nezu J.;
RT "Mouse OCTN1 : Polyspecific organic cation transporter.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AB016257; BAA36626.1;
DR MGI; MGI:1353479; SLC22A4.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 553 AA; 62290 MW; C4D66BC061398653 CRC64;
```

```
Query Match 86.8%; Score 2470; DB 11; Length 553;
Best Local Similarity 84.6%; Pred. No. 6.4e-144;
Matches 468; Conservative 39; Mismatches 44; Indels 2; Gaps 1;

QY 1 MRDYDEVIAFLGEGWGPFORLIFFLLSASIIPIGNGMGSVVFVFLAGTPEHRCRVPDAAANLSS 60
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MRDYDEVIAFLGEGWGPFORLIFFLLSASIIPIGNGMGSVVFVFLAGTPEHRCRVPDVTNLS 60
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 61 AWRNNVPLRLRDGVEPHSCSRYLATIANFSALGLEPGRDVLGQLESCLDGWEYS 120
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

```
Db 61 SWRNHSIPLETKDGRQVPOSCRRYRLATIANFSAMGLEPQDVLEQLESCLDGWEYD 120
QY 121 QDYVLSVTWTEWNLVCEENKWKPLTTSLSFFVGVLLGSFVSQGLSDRFGKKNVLFATMAVQ 180
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 KDVELSTIVTEWNLVCEENKWKPLTTSLSFFVGVLLGSFVSQGLSDRFGKKNVLFATMAVQ 180
QY 181 TGFSEFQIFSIWEMFTVLFVIVGMQISNYVAVFTLGTTEILGKSVRIIFSTLGVCTFFA 240
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 TGFSEFQIFSNWEMFTVLFVIVGMQISNYVAVFTLGTTEILGKSVRIIFSTLGVCTFFA 240
QY 241 VGYMLPLFAFYFIRDRWMLLLALTVPGLVPLWMTIPESPRWLSIORRPREAEDIQKA 300
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 IGYMVLPLFAFYFIRDRWMLLLALTVPGLVPLWMTIPESPRWLSIORRPREAEDIQKA 300
QY 301 AKMNTAVPAVIEDSVE--ELNPLKQOKAFILDLFTRNIAIMTMSLLWMLTSVGYFA 358
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 AKMNSIVAPAGIEDPDELQELNSLKQKVFILDLFTRNIAITVMAVWMLWMLTSVGYFA 360
QY 359 LSLDAPNLHGDAYLNCFLSALIEIPAYITAWMLLRLTPRRYIIAAVLFWGGVLLFIQIV 418
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 361 LSLNPNLHGDAYLNCFLSGLIEVPAYITAWMLLRLTPRRYIIAGVLFWGGVLLFIQIV 420
QY 419 PVDYFISIGLVMKFGITSAFMSLYVFTAEYPTLVNRNAVGVTSASRVGSIITAPYF 478
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 421 PEDYNFYSIGLVMKFGITSAFMSLYVFTAEYPTLVNRNAVGVTSASRVGSIITAPYF 480
QY 479 VYLGAYNRLPYIVMGLSVTLVIGLFTLFPESLGMTLPETLEOMQKVKWFRSGKKTDRSM 538
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 481 VYLGAYNRLPYIVMGLSVTLVIGLFTLFPESLGMTLPETLEOMQKVKWFRSGKKTDRSM 540
QY 539 ETEENPKVLITAF 551
:||||:||||:
Db 541 DREENPKVLITAF 553
```

RESULT 5

```
Q9WTN6 Q9WTN6 PRELIMINARY; PRT; 564 AA.
ID Q9WTN6;
AC Q9WTN6;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DE ORGANIC CATION TRANSPORTER.
GN SLC22A9 OR OCTN3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Nezu J.;
RT "Mouse OCTN3 - a novel OCTN transporter family protein.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AB018436; BAA78343.1;
DR MGI; MGI:1929481; SLC22A9.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 564 AA; 63320 MW; C37FDCA6395DAD01 CRC64;
```

```
Query Match 71.5%; Score 2034.5; DB 11; Length 564;
Best Local Similarity 69.0%; Pred. No. 3.3e-117;
Matches 380; Conservative 75; Mismatches 85; Indels 11; Gaps 3;
```

```
QY 1 MRDYDEVIAFLGEGWGPFORLIFFLLSASIIPIGNGMGSVVFVFLAGTPEHRCRVPDAAANLSS 60
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MRDYDEVIAFLGEGWGPFORLIFFLLSASIIPIGNGMGSVVFVFLAGTPEHRCRVPDVTNLS 60
:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
```

```
QY 61 AWRNNSVPLRLDGRVPHSCSRYLATIANFSALEGRDVLQDLQESCLDGEWFS 120
D 61 AWRNHSIPMETKQGPVQKRRYRLATIANFSELGLEPGRDVLQDLQESCLDGEWYD 120
QY 121 ODVYLSVTVTEWNLVCEDNKKVPLTSLFVGVLLGSFVSGQSLDRFGKKNVLFPATMAVO 180
D 121 KDIFLSTIVTEWDLVCKDDWKAPLTTFFVGVLLGSFVSGQSLDRFGKKNILFTMAH 180
QY 181 TGFSLQIFISWEMTVLFIIVGMQISNYVAVFLLTGKSKVRITFSLTGCTFEA 240
D 181 TGFSLQIFISWEMTVLFIIVGMQISNYVAVFLLTGKSKVRITFSLTGCTFEA 240
QY 241 VGYMLPLFAFYTRDWRMLLALTVPGLVCLPWFIPESPRWLISQRRFAEDIQKA 300
D 241 FGPWLPLFAFYTRDWRMLLALTVPGLVCLPWFIPESPRWLISQRRFAEDIQKA 300
QY 301 AKNNVTAPVAFVDSVEELNPL-----KQKAFILDLFRTNIAIMTMSLLMLTSV 354
D 301 AKINGIVAPSTIFDP-SETNKLQDSSKKPQSHHIDVLRTPNIRITLITMSIILTISV 359
QY 355 GYFALSIDAPNLHGDAYLNCFLSALIEIPAYITAWMLLRLTPRXYIAAVLFGWGGVLLF 414
D 360 GYFGLSLDPLNGLNYVNCFLAAVEPAYVLAWLLQHVSRRTSMAGSLFGSVLL 419
QY 415 IQLVPVDYFLSLTGLVLMKFGITSFMSLYVFTAEYPTLVNMAVGVTSTASRVGSI 474
D 420 VOLVPSDLHLVSTLTMVKGFGITSAYSVMVYVYTAELYPTVVVNNMGVGSSTASRLGSTL 479
QY 475 APVFEVLYGAYNMLPLVIVNGSLTVLIGITLFFPESLGMTLPTLEQMKQKVFGRGKT 534
D 480 SPVFEVLYGAYNMLPLVIVNGSLTVLIGITLFFPESLGMTLPTLEQMKQKVFGRGKT 534
QY 535 RDSMETEENPK 545
D 536 RQSLSKKSPK 546
RESULT 6
Q9VCA2 PRELIMINARY; PRT; 548 AA.
AC Q9VCA2;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 17, Last annotation update)
DE ORCT PROTEIN.
GN ORCT OR CG6331.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
```

```
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner B.C., Riden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AE003747; AAF56271.1;
DR FlyBase; FBgn0019952; Orct.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 548 AA; 61002 MW; 08D7F97599B477AF CRC64;

Query Match 30.1%; Score 857; DB 5; Length 548;
Best Local Similarity 35.7%; Pred. No. 4.8e-45;
Matches 201; Conservative 98; Mismatches 208; Indels 56; Gaps 10;

QY 4 YDEVIAFLGEMGPQRLLFFLLSIIIPNGFNGSVVFLAGTPEHRCRVP-----DAANLS 59
D 3 YDDVITHLGEFGPYKRIYVLLCPALVCAFHKLGLAGVLLAKDFCALPYENGSIYELS 62
QY 60 SAWENNSVPLRLDGRVPHSCSRYLATIANFSALEGRDVL-----COL----- 108
D 63 PHLNLSYPENER-----CSYI-----DVDYTEEYLNCSIPRSSN 97
QY 109 EQESCLDGEWFSQDVLSTVVTWNLVCEDNKKVPLTSLFVGVLLGSFVSGQSLDRFG 168
D 98 ETKTC-SSYVYDRSKYLSNAVTEWNLVCSRLSLSATSLSFMLGVLLGLSFLFCQMSDKLG 156
QY 169 RKNVLPATMAVQGFSPLOIFSLISWEMFTVLFVIVNGQISNYVAVFLLTGKSKVRI 228
D 157 RKPTFFASLVLQIFGVLAAPAPEYFSTISRMIVGATTSVGVFLVAVIALEMGVSSYR- 215
QY 229 IFSTLGVCTFFAVGYMLLPLFAFYTRDWRMLLALTVPGLVCLPWFIPESPRWLISQR 288
D 216 LFAGVAMQMFSSVGFMLTAGFAFYTRDWRMLQTAITLPGLLFLCYWYIIPESARWLLMG 275
QY 289 RFRAREDIIQAAKMMNNTAVPAVIFDS-VEELNPLKQQ-----KAFITDLFRTRNIA 339
D 276 RKDEAFVIEKAANKENVEPNEIYEQVDEVAEKKKQDEMAASQAPATVDFLLRYPNLR 335
QY 340 IMTMSLLMLTISVGFALSIDAPNLHGDAYLNCFLSALIEIPAYITAWMLLRLTPRY 399
D 336 RKTLLIFDFNFVNGVYVGLSNWNTNLGGNQLVNFMSIGAVEIPGYVTLTLLFTLNRRGRS 395
QY 400 ITAAVLFWGGVLLFQLVDPVDYFELSIGLVMGLKFGITSFMSLYVFTAEYPTLVNRM 459
D 396 ILGCTMMVAGISLLATIFVPSDMNMLIVACAMGLKAIATSSYGTITIFSAEQFPVVRNV 455
QY 460 AVGVTSTASRVGSIAPFYVLYGAYNMLPLVIVNGSLTVLIGITLFFPESLGMTLPTL 519
D 456 GLGASSMVARVGGTILAPYLKLLGEIWRPLPIICGALSITAGLLSLLTETLNKMPETI 515
QY 520 EQMKQKVFGRGKTDRSDMETEE 542
D 516 EDGENF-----GKKPAQETAEE 533
```



```
RESULT 7
O01384
ID O01384 PRELIMINARY; PRT; 548 AA.
AC O01384;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE ORGANIC CATION TRANSPORTER.
GN ORCT OR CG6331.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98072431; PubMed=9409773;
RX Taylor C.A.M., Stanley K., Shiras A.D.;
RT "The orct gene of Drosophila melanogaster codes for a putative organic
RT cation transporter with six or 12 transmembrane domains.";
RL Gene 201:69-74(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; Y12400; CAA73031.1;
DR FlyBase; FBgn0019952; Orct.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 548 AA; 61105 MW; 2665B71C969C5E4A CRC64;

Query Match 29.9%; Score 852; DB 5; Length 548;
Best Local Similarity 35.5%; Pred. No. 9.8e-45;
Matches 200; Conservative 99; Mismatches 208; Indels 56; Gaps 10;

QY 4 YDEVIAGFGWGPORLIFELLSSAIIPIGNGMSVVFVLAGTPEHRCRVP-----DAANLS 59
DB 3 YDDVITTHGEFGPKQRIYLLCLPAIVCAFHLKAGVFLAKDFDCAIPYNGSIYELS 62
QY 60 SAWRNNSVPLRLDRGVRPHSCRYRLATIANFSALGLEPGROVDL-----COL----- 108
DB 63 PHLNLSYPENER-----CSYV-----DVDVTEEYLGNSIPRSSN 97
QY 109 EQESCLGWGFEQDYLSTVVTENLVCEDNWKVPLTTSFFVGVLLGSFVSQLSDFEG 168
DB 98 ETKTC-SYIVYDRSKYLSNAVTEWNLVCSRLSATSDSLFGLVGLGSLIFGQMSDKLG 156
QY 169 RKNVLFATMAVOTGFSFLQIFSIWEMFTLVFIVGMQISNVYVAFILGTEILGKSVRI 228
DB 157 RKPTFFASVLQILFGVLAAPVPEYFTTIRMLVGGATTSGVFLVAVIALEWVGSSYR- 215
QY 229 IFSTLGVCTFFAVGYMLPLPFAFIRDMRLLLALTVGVLCVPLMWFIPESRWLISQR 288
DB 216 LFAGVAMQMFPSVGFMLTAGFAYFIDWRWLQIAITLPLGLFLCYWIIIPESARLLMKG 275
QY 289 RPRAEETIQAAKNNTNPAVIFDS-VEELNPLKQ-----KAFILDIFERTNIA 339
DB 276 RKDEAFVIEAKAENKVEVPEIYEQLVDEVAEKKLSODEMAASQAPATVFDLLRVPNLR 335
QY 340 IMTMSLLWMLTSGVYFALSDAPNLHGDAYLNCFLSALIEIETPAYITAMLLRLTPRY 399
DB 336 RTLLIFDFWVNGSVYGLSNWNTNNGNOLVNFMLSGVPEIPGYTLFFTLNRWGRS 395
QY 400 ITAAVLEWGGVLLFQIOLVPDYVDFYLSGLVLMKFGKFTSAFSLMYVFTAEYPTLVNM 459
DB 396 ILCGTMVAVIGLSLLATFVPSDMNWLIVACAMIGKLAITSYGTIYIFSAEQFPVVRNV 455
QY 460 AVGVTSTASRVGSIAPYFVYLGAYNMLPYIVMGSLTVLIGITLFFPESLGMTLPETL 519
DB 456 GIGASSMVARVGGIAPYKLLGLTWIRPLPLIICGALSITAGLLSRLPPTLNKMPETI 515
```

```
QY 520 EQMKVKWFRSGKKTRDSMETEE 542
DB 516 EDGENF-----GKPAPOETAEE 533

RESULT 8
O9VCA3
ID O9VCA3 PRELIMINARY; PRT; 567 AA.
AC O9VCA3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CGI3610 PROTEIN.
GN CGI3610.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Hoch A.S., Dunkov B.C., Dunn P.,
RA Durbin K., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yang Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AE003747; AAF56270.1;
DR FlyBase; FBgn0039176; CGI3610.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 567 AA; 63166 MW; 0D9B979469395E60 CRC64;

Query Match 28.8%; Score 819; DB 5; Length 567;
Best Local Similarity 34.0%; Pred. No. 1.1e-42;
```


459 NWLVVTLFLVGLKLTSSFAVITYTTAEEMPTVIRSGGVGMSTFAFGAMLAPFVPLLA 518
 483 AYNRMLPYIVMGSLTVLIGLFTLPFPEISGLMTPET 518
 519 SYDPLPLLLFGILUSVAGLSLLLPETENRKLPT 554

```

RESULT 10
Q9U539
ID Q9U539 PRELIMINARY; PRT; 568 AA.
AC Q9U539;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE ORGANIC CATION TRANSPORTER 1.
GN OCT-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
[1]
RN R1
RP MEDLINE=99227113; PubMed=10209228;
RA Wu X., Rei Y.J., Huang W., Chancy C., Leibach F.H., Ganapathy V.;
RT "Identity of the P52F12.1 gene product in Caenorhabditis elegans as an
RT organic cation transporter."
RL Biochim. Biophys. Acta 1418:239-244(1999).
RC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DE EMBL: AF110415; AAF21932.1; -.
DR InterPro: IPR003662; sub.transporter.
DR Pfam: PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 568 AA; 63505 MW; COA3E73851F44056 CRC64;

Query Match 26.5%; Score 754.5; DB 5; Length 568;
Best Local Similarity 30.2%; Pred. No. 9,7e-39;
Matches 169; Conservative 119; Mismatches 236; Indels 35; Gaps 9;

```

117	STYLDLSTVEENLVCDOQAAWIEISTSTFYVGSFIGNCLFGYVADKFGRRSRFFVILT	VLI	176
182	GFSPLOFTSISWEMFTVLFWIVGMQISNVVAFILGTELKSVRIIPSTLGVCTFEAV	241	
177	VCGTASSFAKDIESFIRLEFTGLAPFALQIPFIICMEFGMSGR-IFSGMLTSLFGA	235	
242	GYMLLPFLFAYFIRDWRMLLALTVPGVLGVPLWKFIPESPRWLIISORREAEZIIQAA	301	
236	AMALLGVVMEIRKWRQITFCNAPPAFYIIYFFIPESPRNSVSGVKWADAKKQLKTA	295	
302	KMN-----NTAVPAVIFDSVEEENLPLKQOKAF-----ILDJFRTNIAIMTISMLLWMLTS	353	
296	KMGKSNVDVDELV-DSMKNHQNAABEKETKRSHNVTDLFKTLNLRKRLTIIVYI	354	
354	VGYPFALSDAPNLHGDAYLNCFLSALTEIPAYITAWMLLTLPRLRYIIIAVLFWGGVLL	413	
355	IITYNGLTILNVLNPDDYWSRIINGAVELPGYFVVPWPLQOCARRWTLAATMIVCGICV	414	
414	FTQLVPVDYFYLSIGLYMLGRFGTTSAFSMLYVFTABLYPTLVNRNMAVGVTSTASRVGI	473	
415	SAMFPDGYPMLVASATFGRKFGVCSGFVYIIPAGELYFTVYVRAICMGKSSMVACSGLL	474	


```

QY 174 FATMAVQTGFSFLQIFSISWEMFTVLVIVGMGOISNYWAFILGTEILGKSVRIESTL 233
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 LITVLINAAGVLMMAISPTYTWMILFRILQGLVSKAGWLGIVILITEFVGRRYR---RTV 237

QY 234 GYC--TFFAVGYMLPLPAYFIRDRWRMLLLALTVPGLVCLPWFIPESRWLISQRFR 291
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 GIFYOVAYTVGLLVAGVAYALPHRWLQFTVALPNFFELLYWYCIPESRWLISQNKNA 297

QY 292 EADIIIOAKAKNNNTAVPAVI-----PDSVEELNPKQAKFILDLPFRTRNIAIMTMS 345
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 298 EAMRIIKHTAKNGSLSPASLQRLLEETGKKLNP-----SFDLVRTPOIRKHTMIL 351

QY 346 LLLWMLTSVGYFALSADPNLHGD-AYLNCFLSALIEIPAYITAMLLRLTPRRYIIAAV 404
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 352 MYNWFSSVLYOGLNHHM-GLAGDNIYLDFFYSALVEPFAAPMILITIDRIGRRYPWAAS 410

QY 405 LFWGGGVLLFIQLVPVDYFSLGVLMLGKFGITSAFSMLYVFTAEYLTPTLVIRNMAVGT 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 411 NMVAGAACLASVFIQDGLQWLKIIISCLGRMGITMAYEIVCLVNAELYPTETIRNLGVHIC 470

QY 465 STASRVGSIIPFYVYLGAYNRM-----LPYIVMGSLTVLIGIETLFPESLGMTPET 518
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 471 SSMCDIGGIITPFLY-----RLTNWLEPLMVFVGLVAGGLVLLLPETKGALEPET 525

QY 519 LEOMQKWKFRSGKK 533
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 526 IEAEANQRPKNKE 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
O02713 PRELIMINARY; PRT; 554 AA.
AC O02713;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE APICAL OPTICAL CATION TRANSPORTER.
GN OCT-2p.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Sulina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97256752; PubMed=9099681;
RA Gruendemann D., Babin-Ebell J., Martel F., Oerding N., Schmidt A.,
RA Schoemig E.;
RT "Primary structure and functional expression of the apical organic
RT cation transporter from kidney epithelial LLC-PK1 cells.";
RL J. Biol. Chem. 272:10408-10413(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; Y09400; CAA70567.1; -.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar tr. 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 554 AA; 61989 MW; E14B5565600C553B CRC64;

```

Query Match 24.6%; Score 700; DB 6; Length 554;
 Best Local Similarity 34.0%; Pred. No. 2.1e-35;
 Matches 188; Conservative 94; Mismatches 237; Indels 34; Gaps 13;

```

QY 1 MRDYDEVIAFLGEGWPPQRIFF---LLSASIIIPNGFNGSVVFLAGTPEHRCRVPDAAN 57
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MLTVDDILEHGEFNFQKQTFLLALLSAFTAFT---IYVGIVFLGFIPIHRCRSPGVAE 57

QY 58 LSS--ANR-----NNSVPLRLRDGREVPHSCSRYLATIANFSAIG-LKP--GRDVLGQ 107
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 58 LSQRCGWSLAELNLTVPGGPAGQAFPROCRRYEVDW---NQSLGCVDPPLAGLAANSSH 115

```

```

QY 108 LEQESCLDGEFSDQVYLVSTVTWNLCEDNNKVPPLTTSLFEVGVLLGSFVSGQUSDRE 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 LPLGPCRYGWY--DTPGSSIVTEFDLVCANSWLLDLFQSAVNVGFFIGSVGIGYIADRF 173

QY 168 GRKNVLPAFMVQTGFSFLQIFSISWEMFTVLVIVGMGOISNYWAFILGTEILGKSVR 227
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 174 GRKICLLLTILINAVSGVLMMAISPTYTWMILFRILQGLVSKAGWLGIVILITEFVGLSYR 233

QY 228 IIFSTLQVC--TFFAVGYMLPLPAYFIRDRWRMLLLALTVPGLVCLPWFIPESRWLI 285
   || : || : || : || : || : || : || : || : || : || : || : || : ||
Db 234 ---RTVGIQVQVAFTFGLLVLAGVAYALPHRWLQFTVTLPNFCFLFYWCVPESRWLI 290

QY 286 SQRREFREADIIOAKAKNNNTAVPAVIFDSVEELNPKQ----OKAFILDLFRRNIAIM 341
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 291 SONKNAKAMSIIKHIANKNGKSLPA-----SIQSLRPDEEVEGKLPKSFLDLVRTPQIRKH 346

QY 342 TIMSLLLWMLTSVGYFALSADPNLHGDAYLNCFLSALIEIPAYITAMLLRLTPRYII 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 347 TLILMYNWFSSVLYOGLVHMHGSLAGNLYLDFYSALVEPFAAPLILLLTIDRLGRHPW 406

QY 402 AAVLFWGGGVLLFIQLVPVDYFSLGVLMLGKFGITSAFSMLYVFTAEYLTPTLVIRNMAV 461
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 407 AASNVVAGAACLASVFIQDGLQWLKIIISCLGRMGITMAYEIVCLVNAELYPTETIRNLGV 466

QY 462 GVTSTASRVGSIIPFYVYLGAYNRMPLPYIVMGSLTVLIGIETLFPESLGMTPETLE 520
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 467 LVCSSMCDIGGIITPFLYRLDIWHELPLVFAVAVGLIAGGLVLLLPETKGTLPETIE 526

QY 521 OMQKWKFRSGKK 533
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 527 EAETMRPRKNKE 539
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
P70485 PRELIMINARY; PRT; 593 AA.
AC P70485;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCT2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
RX MEDLINE=96295517; PubMed=8702418;
RA Okuda M., Saito H., Urakami Y., Takano M., Inui K.;
RT "cDNA cloning and functional expression of a novel rat kidney organic
RT cation transporter, OCT2.";
RL Biochem. Biophys. Res. Commun. 224:500-507(1996).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; D83044; BAA11754.1; -.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar tr. 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 593 AA; 66080 MW; 36C1044E0C04B443 CRC64;

```

Query Match 24.6%; Score 699.5; DB 11; Length 593;
 Best Local Similarity 34.0%; Pred. No. 2.4e-35;
 Matches 189; Conservative 87; Mismatches 243; Indels 37; Gaps 14;

```

QY 1 MRDYDEVIAFLGEGWPPQRIFF---LLSASIIIPNGFNG--MSVVFVLAGTPEHRCRVPDAANL 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MSTVDDILEHGEFNFQKQTFLL--ALLSGAFTPIYVGIVFLGFTPDHHCWSPGAACL 58

QY 59 SS--AWR-----NNSVP-LRLRDGREVPHSCSRYR-----LATIANFSAIGLEPGRDV 103
   || : || : || : || : || : || : || : || : || : || : || : || : ||

```

Db 59 SQRGWSQAELNNTVPLGPSDEASFLSQCMRYEVDNQSTLDCVDPSSLA-----A 112
QY 104 DLGLEQESCLDGEFSQDYLSTVVTENNLCEDNNKYPLTSLTFVGVILGSEVSGOL 163
Db 113 DRNQLPLGCEHGWYNTFG--SSIVTEFNLCAHSWMLDLFQSVVNVGFFIGAMMIGYL 170
QY 164 SDRGRKNVLFAPMAYOTGFSFLQIFSISWEMFTVLFIIVGMGOISNYVVAFILGTEILG 223
Db 171 ADRFGRKFCLLVTILINAISGALMAISPNYANMLVFRFLQGLVSKAGWILIGYLITEFVG 230
QY 224 KSVRIIFSLGVC-LTFPVGVMLLPLPAYFIRDRMLLLALTVPGLVCPVPLWFIPESP 281
Db 231 LGYR---RMVGICYQTAFTVGLLILAGVAYIPNNRWLQFAVTLFPNFCFLLYFNCIPESP 287
QY 282 RWLISQRREREADIIQKAAKNNTAVPAVIFDSVEELNPLAQKQAFILDLFRTNRNIAIM 341
Db 288 RWLISQNKIVKAMKIKKHAKKNGKSVPSLQNLTPDEDAKGLKPSILDLVRTPOIRKH 347
QY 342 TMSLLMLMTSVGYFALSADAPNLHGD-AYLNCFLSALIEIPAYITAWLLRLTLPRRYI 400
Db 348 TLILMYNFTSSVLYQGLIMHM-GLAGDNIYLDFFYSALVEPPAAFIILITIDRVGRYP 406
QY 401 IAAVLFWGGGVLLFIQLVPDYFFLSIGLVMLKFGITSAFSMLYVFTAELYPTLVNRMA 460
Db 407 WAVSNMVAGAACLASVFIPTDDQLWLKITTACLGMRGITMAYEMVCLVNAELYPTVIRNLG 466
QY 461 VGYTSTASRVGSIIAPFYVY-LGAYNRMLPYIVMGSLTVLIGIFTLPFPESIGMTLPETL 519
Db 467 VLVCSMCDIGGIIIPFLVYRLTDIWMEEPLVVFVAVGLVAGALVLLLPETKGPALPETI 526
QY 520 EQMKVKWFRSGKKTR 535
Db 527 EDAENMQ--RPRKKER 540

Search completed: January 25, 2002, 10:14:54
Job time: 335 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:09:12 ; Search time 43.11 Seconds

(without alignments)
957.059 Million cell updates/sec

Title: US-09-521-195-3

Perfect score: 2883

Sequence: 1 MRDYDEVTAFLGEGWPFQRL.....HTRMLKDQERPIILKSTAF 557

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2883	100.0	557	20	AAV01650
2	2883	100.0	557	21	AAV83929
3	2513	87.2	557	20	AAV01652
4	2513	87.2	557	21	AAV83930
5	2513	87.2	557	21	AAV83930
6	2295.5	79.6	564	21	AAV83930
7	2218	76.9	551	20	AAV01649
8	2148	74.5	553	20	AAV01651
9	2148	74.5	553	21	AAV83930
10	785.5	27.2	560	21	AAV83930
11	784	27.2	584	22	AAV00930

12	735	25.5	554	19	AAW64538	Human liver cell c
13	724.5	25.1	540	22	AAV84901	Murine organic ani
14	719	24.9	535	21	AAV51249	Rat liver anion tr
15	712	24.7	536	17	AAV77676	Rat OCT-1 protein.
16	701.5	24.3	538	22	AAV47273	HOAT2B. Homo sapi
17	697	24.2	546	22	AAV47272	HOAT2A. Homo sapi
18	693	24.0	542	21	AAV92902	Human cerebral org
19	690	23.9	548	21	AAV08823	A human organic an
20	689	23.9	542	22	AAV47274	HOAT3.. Homo sapie
21	686.5	23.8	551	21	AAV08824	A human organic an
22	673	23.3	561	18	AAV44196	Human osteoclast t
23	668	23.2	578	22	AAE06571	Human protein havi
24	660.5	22.9	607	21	AAV12131	Hydrophobic domain
25	656.5	22.8	536	21	AAV92903	Rat cerebral organ
26	652.5	22.6	537	18	AAV44195	Mouse osteoclast t
27	644	22.3	545	22	AAV36553	Mouse organic anio
28	641	22.2	550	21	AAV44278	Human organic anio
29	641	22.2	550	22	AAV47271	HOAT1.. Homo sapie
30	638.5	22.1	551	20	AAV88488	Rat organic anion
31	628	21.8	563	20	AAV88489	Human organic anio
32	621.5	21.6	483	22	AAV00982	Human bone marrow
33	607	21.1	550	22	AAE06612	Human protein havi
34	607	21.1	550	22	AAV89091	Human organic anio
35	541	18.8	541	22	AAV47276	HOAT5. Homo sapie
36	477.5	16.6	553	22	AAE04897	Human transporter
37	451	15.6	554	22	AAV47275	HOAT4. Homo sapie
38	433	15.0	480	21	AAV28044	Arabidopsis thalia
39	433	15.0	483	21	AAV28043	Arabidopsis thalia
40	380.5	13.2	515	21	AAV42522	Arabidopsis thalia
41	368.5	12.8	521	21	AAV30885	Arabidopsis thalia
42	357.5	12.4	548	21	AAV44633	Human organic cati
43	344.5	11.9	397	21	AAV30886	Arabidopsis thalia
44	332	11.5	439	22	AAV76766	Corynebacterium gl
45	332	11.5	448	22	AAV93041	C glutamicum prote

ALIGNMENTS

RESULT 1
AAV01650
ID AAY01650 standard; Protein; 557 AA.
XX
AC AAY01650;
XX
DT 23-JUN-1999 (first entry)
XX
DE A protein with cation transporting activity.
XX
KW Organic cation transporter; OCT1; OCT2; drug development; fatty liver;
KW heart disease; cancer; anti-tumour drug; anticancer drug.
XX
OS Homo sapiens.
XX
PN W09913072-Al.
XX
PD 18-MAR-1999.
XX
PF 07-SEP-1998; 98WO-JP04009.
XX
PR 20-MAY-1998; 98JP-0156660.
PR 08-SEP-1997; 97JP-0260972.
XX
(CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
Nezu J, Oku A;
XX
WPI; 1999-215062/18.
XX
N-PSDB; AAX26880.
XX
Genes homologous with organic cation transporters OCT1 and OCT2,
XX
useful in design of new drugs for treatment of diseases due to
XX
abnormality of the transporter functions

XX PS Claim 1; Page 51-55; 97pp; Japanese.

CC The present sequence represents a protein with cation transporting activity. The genes are significantly homologous with organic cation transporters OCT1 and OCT2. The genes may be used in drug development, particularly in the treatment of diseases due to abnormality of the organic cation transporter functions e.g. fatty liver, heart diseases and cancers, by controlling such as by inhibition or activation.

CC Administration of anti-tumour and anticancer drugs in combination with a transporter protein inhibiting agent allows the agents to penetrate into the diseased cells to enhance the drug action.

XX SQ Sequence 557 AA;

Query Match 100.0%; Score 2883; DB 20; Length 557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRDYDEVTAFLGEGWGPQRLIFLLSASITIPNGFTGLSSVFLIATPEHRCRVPDAAANLSS 60
DB 1 mrdydevtaflgewgpqrlifllsasiipngftglssvfliatpehrcrvpdaanlss 60

QY 61 AWRNHTVPLRLDRGVRPHSCRRYRLATIANFSALEPGRDVDLQLEQESCLDQWEFS 120
DB 61 awrnhtvplrlldrgrevphscrryrlatiansaalepgrdvdlqleqescldgwefs 120

QY 121 QDYLSTIVTEWNLVCEDDWKAPLTISLFFVGVLLGSFISGQLSDRFGKKNVLFVTMGQ 180
DB 121 qdylstivtewnlvceddwkaptislsffvgvllgsgqlsdrfgrknvlfvtmgmq 180

QY 181 TGFSELIQIFSKNFEMFVFLVGVGMGOISNYAAFLVGLTEILGKSVRIIFSTLGVCFYA 240
DB 181 tgfseqliqifsknfemfvlvlgvgmgoisnyaaflvglteilgksvriifstlgvcfiya 240

QY 241 FGVMVPLPFAFIRDRWMLLVALTMPGVLCVALWVFIPESPRWLISQGRFEEAEVIRKA 300
DB 241 fgvmvplpfaafirdrwmlvaltmpgvlcvalwvfipesprwlisqgrfeaeaeiirka 300

QY 301 AKANGIVVPSTIFDPSLQDLSKKQSHNILDLLRTNIRMTIMSLMWTISVGYFG 360
DB 301 akangivvpstifdpselqdlsskkqshnildllrtwnirmvtimsmlmwtisvgyfg 360

QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLILQYLPRIYSMATALPLGGSVLLFMQLV 420
DB 361 lsldtpnlhgdifvncflsamvevpayvlawlilqylpriysmatallpggsvllfmqlv 420

QY 421 PPDLYYLATVLMVGMKFGVTAAFSMVYVYTAELPTVTVVRNMGVSVSTASRLGSILSPYF 480
DB 421 ppdlYYLATVLMVGMKFGVTAAFSMVYVYTAELPTVTVVRNMGVSVSTASRLGSILSPYF 480

QY 481 VYLGAYDRFLPYLLMGSLTILTAITLFLPESFGPLPDTIDQMLRVKGMKHKRTPSHR 540
DB 481 vylgaydrflpyllmgsltiltaltlflpesfgplpdtidqmlrvkgmkhkrtpshtr 540

RESULT 2
AA183929
ID AA183929 standard; protein; 557 AA.
XX
AC AA183929;
XX
DT 05-JUL-2000 (first entry)
XX
DE Human carnitine transporter protein OCTN2.
XX
KW Organic cation transportation; human; carnitine transporter protein;
KW OCTN2; diagnosis; systemic carnitine deficiency; mutation; gene therapy;

KW OS juvenile visceral steatosis.

XX Homo sapiens.

XX WO200014210-A1.

XX PD 16-MAR-2000.

XX PF 07-SEP-1999; 99WO-JP04853.

XX PR 07-SEP-1998; 98JP-0252683.

XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.

XX PI Nezu J, Oku A;

XX WIPI; 2000-256966/22.

XX DR N-PSDB; AAA09889.

XX PT Systemic carnitine deficiency gene OCTN2 encoding part of organic cation transporter, useful as diagnostic tool -

XX PS Example 1; Page 36-40; 106pp; Japanese.

XX CC This sequence represents the human carnitine transporter protein OCTN2. The coding sequence can be used as a target for diagnosis of systemic carnitine deficiency by detecting the presence of mutations in the sequence, especially seen in the disease juvenile visceral steatosis (jvs). The wild type OCTN2 gene can be used in the gene therapy of the disease state.

XX SQ Sequence 557 AA;

Query Match 100.0%; Score 2883; DB 21; Length 557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRDYDEVTAFLGEGWGPQRLIFLLSASITIPNGFTGLSSVFLIATPEHRCRVPDAAANLSS 60
DB 1 mrdydevtaflgewgpqrlifllsasiipngftglssvfliatpehrcrvpdaanlss 60

QY 61 AWRNHTVPLRLDRGVRPHSCRRYRLATIANFSALEPGRDVDLQLEQESCLDQWEFS 120
DB 61 awrnhtvplrlldrgrevphscrryrlatiansaalepgrdvdlqleqescldgwefs 120

QY 121 QDYLSTIVTEWNLVCEDDWKAPLTISLFFVGVLLGSFISGQLSDRFGKKNVLFVTMGQ 180
DB 121 qdylstivtewnlvceddwkaptislsffvgvllgsgqlsdrfgrknvlfvtmgmq 180

QY 181 TGFSELIQIFSKNFEMFVFLVGVGMGOISNYAAFLVGLTEILGKSVRIIFSTLGVCFYA 240
DB 181 tgfseqliqifsknfemfvlvlgvgmgoisnyaaflvglteilgksvriifstlgvcfiya 240

QY 241 FGVMVPLPFAFIRDRWMLLVALTMPGVLCVALWVFIPESPRWLISQGRFEEAEVIRKA 300
DB 241 fgvmvplpfaafirdrwmlvaltmpgvlcvalwvfipesprwlisqgrfeaeaeiirka 300

QY 301 AKANGIVVPSTIFDPSLQDLSKKQSHNILDLLRTNIRMTIMSLMWTISVGYFG 360
DB 301 akangivvpstifdpselqdlsskkqshnildllrtwnirmvtimsmlmwtisvgyfg 360

QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLILQYLPRIYSMATALPLGGSVLLFMQLV 420
DB 361 lsldtpnlhgdifvncflsamvevpayvlawlilqylpriysmatallpggsvllfmqlv 420

QY 421 PPDLYYLATVLMVGMKFGVTAAFSMVYVYTAELPTVTVVRNMGVSVSTASRLGSILSPYF 480
DB 421 ppdlYYLATVLMVGMKFGVTAAFSMVYVYTAELPTVTVVRNMGVSVSTASRLGSILSPYF 480

QY 481 VYLGAYDRFLPYLLMGSLTILTAITLFLPESFGPLPDTIDQMLRVKGMKHKRTPSHR 540
DB 481 vylgaydrflpyllmgsltiltaltlflpesfgplpdtidqmlrvkgmkhkrtpshtr 540


```

QY 541 MLKDGQERPTILKSTAF 557
Db 541 mlkdqgerptilkstaf 557

RESULT 3
ID AAY01652 standard; Protein; 557 AA.
AC AAY01652;
XX
XX 23-JUN-1999 (first entry)
DE A protein with cation transporting activity.
XX Organic cation transporter; OCT1; OCT2; drug development; fatty liver;
KW heart disease; cancer; anti-tumour drug; anticancer drug.
XX
XX Mus musculus.
XX
XX W09913072-A1.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98WO-JP04009.
XX
XX 20-MAY-1998; 98JP-0156660.
XX 08-SEP-1997; 97JP-0260972.
XX
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
XX Nezu J, Oku A;
XX
XX WPI; 1999-215062/18.
XX N-PSDB; AAX26902.
XX
XX Genes homologous with organic cation transporters OCT1 and OCT2,
XX useful in design of new drugs for treatment of diseases due to
XX abnormality of the transporter functions
XX
XX Claim 1; Page 75-79; 97pp; Japanese.
XX
XX The present sequence represents a protein with cation transporting
XX activity. The genes are significantly homologous with organic cation
XX transporters OCT1 and OCT2. The genes may used in drug development,
XX particularly in the treatment of diseases due to abnormality of the
XX organic cation transporter functions e.g. fatty liver, heart diseases
XX and cancers, by controlling such as by inhibition or activation.
XX Administration of anti-tumour and anticancer drugs in combination with
XX a transporter protein inhibiting agent allows the agents to penetrate
XX into the diseased cells to enhance the drug action.
XX
XX Sequence 557 AA;
SQ

Query Match 87.28; Score 2513; DB 20; Length 557;
Best Local Similarity 85.58; Pred. No. 2.5e-265;
Matches 476; Conservative 39; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRDYDEVTAFLGEGPFORLIFLILSLASIIIPNGFTGLSSVFLIATPSHRCRVDPDAANLSS 60
Db 1 mrdydevtaflgwgppqrliflilsilpnpgfngmsivflactpchrclvphbtvnlss 60
QY 61 AWRNHTVPLRLDRGVEPHSCRRYRLATIANFSAIGLEPGRDVDLGOLEQESCLDGEWFS 120
Db 61 awrnhsipletkdrqvpqkrryrlatianselgiepgrdvldleqescldgweyd 120
QY 121 QDVYLSTIVTWNLVCEWDKAPITISLFFVGVLLGSFISGQLSDRGRKNVLFVTMGMQ 180
Db 121 kdvlstivtewdlvckdwkapititslffvgvlmgfsglsdrgrknvlfvtmgmq 180
QY 181 TCFSEFLOIFSKNFEMFVVLFLVVGMGQISNYAAAFVLGLTGLKSVRIITSLTGLVCIFYA 240

```

```

Db 181 tgfsflqvsvnfemftvflvlgmgqisnyaaafvlgteillsksiriifatlgvcifya 240
QY 241 FGVMVPLPAYFIRDRMLLVALTMPGVLCVALWFWFPESPRWLISOGRFEAEVIRKA 300
Db 241 fgmvlplfayfirdwrmlilaltvpgvlgcalwfwfipesprwlisqgrikeaeviirka 300
QY 301 AKANGIVVPSITFDPSQLDSSKKQSHNLDLRTWNIRMTIMSTIMLWMTISVGVEFG 360
Db 301 akingivapstifdpselqldinstkpkqlhhydiirtnirvitimsiilwtisvgyfg 360
QY 361 LSLDTPNLHGDIYVNCFLSAMVEYPAYVLAWLLQYLPRRYSMATATFLGGSVLLFMQLV 420
Db 361 lsltdpnlhgdlyvncfllaavepayvlawllqylprrysaalfggsvllfmqlv 420
QY 421 PPDLYLATLVLMVGKFGVTAAFSMAVYVYTAELPYTVVRNMGVGVSSASTASLGSLSPYF 480
Db 421 pslfylstalyvmvgkfgitaysmvyvyytaelyptvvrnmgvgvssastaslgslspyf 480
QY 481 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGPLPDTIDQMLRVKGMKHKRTPSHTR 540
Db 481 vylgaydrflpyilmgsltiltailitlffpesfgvplpdtidqmlrvkgikwqigsqtr 540
QY 541 MLKDGQERPTILKSTAF 557
Db 541 mlkdqgeesptvlkstaf 557

RESULT 4
ID AAB20580 standard; Protein; 557 AA.
XX
XX AC AAB20580;
XX
XX DT 11-DEC-2000 (first entry)
XX
XX DE Mouse OCTN2 amino acid sequence.
XX
XX KW Mouse; transporter; OCTN1; OCTN2; OCTN3; organic cation transporter;
XX identification; regulator; carnitine transport.
XX
XX OS Mus musculus.
XX
XX PN W0200046368-A1.
XX
XX PD 10-AUG-2000.
XX
XX PF 04-FEB-2000; 2000WO-JP00619.
XX
XX PR 05-FEB-1999; 99JP-0028406.
XX
XX PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
XX PI Nezu J, Ose A;
XX
XX DR WPI; 2000-586982/55.
XX
XX PT Organic cation transporter gene OCTN3 expressed in testis for
XX identification of regulators of carnitine transport for use as drugs
XX
XX PS Example 2; Fig 2; 58pp; Japanese.
XX
XX CC The present invention describes a mouse organic cation transporter
XX protein (OCTN3). Also described are: (1) a method for screening
XX compounds for their ability to regulate the transport of an organic
XX cation into the cell, by generating a cell expressing OCTN3 at the
XX cell membrane, contacting with the compound and organic cation, and
XX observing the degree of transport of the organic cation; and (2) a
XX method for screening compounds for their ability to be transported into
XX the cell by OCTN3, by generating a cell expressing OCTN3 at the cell
XX membrane, contacting with the compound and observing the degree of
XX transport of the compound. OCTN3 can be used for the identification of
XX regulators of the transport of organic cations (especially carnitine)

```

CC into cells by OCTN3, for use as drugs. The present sequence represents
 CC the mouse OCTN2 amino acid sequence, which is used in an example from
 CC the present invention.

XX Sequence 557 AA;

Query Match 87.2%; Score 2513; DB 21; Length 557;
 Best Local Similarity 85.5%; Pred. No. 2.5e-265;
 Matches 476; Conservative 39; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRDYDEVTAFLGEGWGPQRLIFFLSASIIIPNGFTGLSSVFLIATPEHRCRVPDPAANLSS 60
 Db 1 mrdydevtaflgwgppqrlifflsasiipngftglssvfliatpehrcrlvphvtvnls 60
 QY 61 AWRNHTVPLRLDGRVPHSCRRYRIATIANFSALEGPGRVDLQLEQESCLDGWEFS 120
 Db 61 awrnhsipletkdgrvphscrryriatianfseglepgrvdleqesclgdgweyd 120
 QY 121 QDYLSTIVTEWNLVCEDDWKAPLTISLFFVGVLLGSFISGQSLDRFGKKNVLFVTMGMQ 180
 Db 121 kvflstivtewdlvcddwkaptltslffvglmgsfsgslsdrfgrknvflvtmgmq 180
 QY 181 TGFSLQIFSKNFEMFVLVVGMCQISNYVAAFVLGTEILGKSVRIIFSTLGVCFIYA 240
 Db 181 tgfsflqvsvnfemftvlvvgmqisnyvaafvlgteilsksvriifatlvgvcifya 240
 QY 241 FGVMVLPFLFAYFTRDWRMLLVALTMPCVLCVALWFWIPESPRWLSQGRFEAEVIRKA 300
 Db 241 fgvmvlplfayftrdwrmlaltvpgvlgcgalwfwipesprwlsqgrikeaeavlirka 300
 QY 301 AKANGIWPSTIFDPSLQDSSKKQOSHNLIDLLRTNIRMTIMSIMLWMTISVGYFG 360
 Db 301 akingiwapstifdpselqdnstkpqlhhydlrtnirvmtimsilwltisvgyfg 360
 QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOLQVLPDPRYSMATALFLGGSVLLFMQLV 420
 Db 361 lsldtpnlhgdlyvncflaavevpayvawlllqylprysisaaflggsvllfmqlv 420
 QY 421 PPOLYLATVLMVGKFGVTAAFSMVYVYTAELYPTVVRNMGVSVSTASRLGSLSPYF 480
 Db 421 pselfylatalvmvgkfgitsaysmvvytaelyptvvrnmvgvsvstasrlgsilspyf 480
 QY 481 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGPLPDTIDQMLRVKGMKHKRTPSHTR 540
 Db 481 vylgaydrflpyilmgsltiltailtflfespfgvplpdtidqmlrvkgikwqigsqtr 540
 QY 541 MLKDGQERPTILKSTAF 557
 Db 541 mqkdgeesptvikstaf 557

RESULT 5
 AAY83930
 ID AAY83930 standard; Protein; 557 AA.
 XX
 AC AAY83930;
 XX
 DT 05-JUL-2000 (first entry)
 XX
 DE Mouse carnitine transporter protein OCTN2.
 XX

KW Organic cation transportation; mouse; carnitine transporter protein;
 KW OCTN2; diagnosis; systemic carnitine deficiency; mutation; gene therapy;
 KW juvenile visceral steatosis.

OS Mus musculus.
 XX
 FN WO200014210-A1.
 XX
 PD 16-MAR-2000.
 XX
 PF 07-SEP-1999; 99WO-JP04853.

XX 07-SEP-1998; 98JP-0252683.
 XX (CHUG-) CHUGAI, RES INST MOLECULAR MEDICINE INC.
 XX PA
 XX PI Nezu J, Oku A;
 XX
 XX WPI; 2000-256966/22.
 DR N-PSDB; AAA09890.
 XX
 PT Systemic carnitine deficiency gene OCTN2 encoding part of organic
 PT cation transporter, useful as diagnostic tool -
 XX
 PS Example 1; Page 47-51; 106pp; Japanese.

CC This sequence represents the mouse carnitine transporter protein OCTN2.
 CC The coding sequence of the corresponding human protein can be used as a
 CC target for diagnosis of the systemic carnitine deficiency by detecting the
 CC presence of mutations in the sequence, especially seen in the disease
 CC juvenile visceral steatosis (jvs). The wild type OCTN2 gene can be used
 CC in the gene therapy of the disease state.

SQ Sequence 557 AA;

Query Match 87.2%; Score 2513; DB 21; Length 557;
 Best Local Similarity 85.5%; Pred. No. 2.5e-265;
 Matches 476; Conservative 39; Mismatches 42; Indels 0; Gaps 0;

QY 1 MRDYDEVTAFLGEGWGPQRLIFFLSASIIIPNGFTGLSSVFLIATPEHRCRVPDPAANLSS 60
 Db 1 mrdydevtaflgwgppqrlifflsasiipngftglssvfliatpehrcrlvphvtvnls 60
 QY 61 AWRNHTVPLRLDGRVPHSCRRYRIATIANFSALEGPGRVDLQLEQESCLDGWEFS 120
 Db 61 awrnhsipletkdgrvphscrryriatianfseglepgrvdleqesclgdgweyd 120
 QY 121 QDYLSTIVTEWNLVCEDDWKAPLTISLFFVGVLLGSFISGQSLDRFGKKNVLFVTMGMQ 180
 Db 121 kvflstivtewdlvcddwkaptltslffvglmgsfsgslsdrfgrknvflvtmgmq 180
 QY 181 TGFSLQIFSKNFEMFVLVVGMCQISNYVAAFVLGTEILGKSVRIIFSTLGVCFIYA 240
 Db 181 tgfsflqvsvnfemftvlvvgmqisnyvaafvlgteilsksvriifatlvgvcifya 240
 QY 241 FGVMVLPFLFAYFTRDWRMLLVALTMPCVLCVALWFWIPESPRWLSQGRFEAEVIRKA 300
 Db 241 fgvmvlplfayftrdwrmlaltvpgvlgcgalwfwipesprwlsqgrikeaeavlirka 300
 QY 301 AKANGIWPSTIFDPSLQDSSKKQOSHNLIDLLRTNIRMTIMSIMLWMTISVGYFG 360
 Db 301 akingiwapstifdpselqdnstkpqlhhydlrtnirvmtimsilwltisvgyfg 360
 QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOLQVLPDPRYSMATALFLGGSVLLFMQLV 420
 Db 361 lsldtpnlhgdlyvncflaavevpayvawlllqylprysisaaflggsvllfmqlv 420
 QY 421 PPOLYLATVLMVGKFGVTAAFSMVYVYTAELYPTVVRNMGVSVSTASRLGSLSPYF 480
 Db 421 pselfylatalvmvgkfgitsaysmvvytaelyptvvrnmvgvsvstasrlgsilspyf 480
 QY 481 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGPLPDTIDQMLRVKGMKHKRTPSHTR 540
 Db 481 vylgaydrflpyilmgsltiltailtflfespfgvplpdtidqmlrvkgikwqigsqtr 540
 QY 541 MLKDGQERPTILKSTAF 557
 Db 541 mqkdgeesptvikstaf 557

RESULT 6
 AAB20578
 ID AAB20578 standard; Protein; 564 AA.

XX AAB20578;
 XX
 XX
 XX 11-DEC-2000 (first entry)
 XX
 XX Mouse OCTN3 protein SEQ ID NO:1.
 XX
 XX Mouse; transporter; OCTN1; OCTN2; OCTN3; organic cation transporter;
 KW Identification; regulator; carnitine transport.
 XX
 XX Mus musculus.
 XX
 XX WO200046368-A1.
 XX 10-AUG-2000.
 XX
 XX 04-FEB-2000; 2000WO-JP00619.
 XX
 XX 05-FEB-1999; 99JP-0028406.
 XX
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX
 XX Nezu J, Ose A;
 XX
 XX WPI: 2000-586982/55.
 XX N-PSDB; AAA88053.
 XX
 XX Organic cation transporter gene OCTN3 expressed in testis for
 PT identification of regulators of carnitine transport for use as drugs -
 PT
 XX Claim 1; Page 34-39; 58pp; Japanese.
 XX
 XX The present invention describes a mouse organic cation transporter
 CC protein (OCTN3). Also described are: (1) a method for screening
 CC compounds for their ability to regulate the transport of an organic
 CC cation into the cell, by generating a cell expressing OCTN3 at the
 CC cell membrane, contacting with the compound and organic cation, and
 CC observing the degree of transport of the organic cation; and (2) a
 CC method for screening compounds for their ability to be transported into
 CC the cell by OCTN3, by generating a cell expressing OCTN3 at the cell
 CC membrane, contacting with the compound and observing the degree of
 CC transport of the compound. OCTN3 can be used for the identification of
 CC regulators of the transport of organic cations (especially carnitine)
 CC into cells by OCTN3, for use as drugs. The present sequence represents
 CC mouse OCTN3.
 XX
 XX Sequence 564 AA;
 XX
 XX
 XX Query Match 79.6%; Score 2295.5; DB 21; Length 564;
 XX Best Local Similarity 80.6%; Pred. No. 1.4e-241;
 XX Matches 435; Conservative 54; Mismatches 48; Indels 3; Gaps 1;
 XX
 XX 1 MRDYDEVTAFLGEMGPFORLFFLLSIIIPNGFTGLSSVFLIATPEHRCRVPDAANLSS 60
 DB 1 mldydevtaflgwtqfqlifllsaiipngftglssvflaipaehrcrripdvnlss 60
 XX
 XX 61 AWRNUTVPLRLDREVPCHSCRRRLATIANFSALEPGRDVLDGQESCLDGEFFS 120
 DB 61 awrnhsipmetkdgevpqkerryrlatianfselepgrdvldgeqclndgweyd 120
 XX
 XX 121 ODVYLSITVTENLVCEDDWKAPITISFFVGLLGSFISQSLDRFGKRVNLFVTGMQ 180
 DB 121 odvylstvtewnlvckddwkapttsffvgvllgsfsgqslsdrfgkvnlfvtnamh 180
 XX
 XX 181 TGFSEFLQIFSKNFEMFVFLVGLVGMGINSYVAAFVLGTEILGKSVRIIFSTLGVCFIYA 240
 DB 181 tgfsvflqifsvnfemftllylvgmghisnyvaafvlgtemlksksvriifatlvgvciffa 240
 XX
 XX 241 FGYMVLPLFAFIRDRMMLVALMPGVLVALWVFIPESPRWLLSQGRFEEAEVILIRKA 300
 DB 241 fgmvlplfayfirewrrllaitlpgvlcalwvfipesprwllsqgrfkeaeavlirka 300

QY 301 AKANGIVVPSTIFDPSE---LQDLSSKKQSHNILDLLRTWNRIMVMTSIMLWMTISVG 357
 DB 301 akingivapstifdpsetnklqddsskpkqshhiydvrtpniriltimsilwltisvg 360
 QY 358 YFGLSLDTPNLHGDI FVNCFLSAMVEVPAYVLAWLLQYLPFRYSMATALFLGGSVLLEFM 417
 DB 361 yfglsldtpnlngniyvcflaavevpayvlawllqhvsvrrysmagsiflg9svilly 420
 QY 418 OLVPDLYLATVLMVGKFGVTAAFSMVYVYTAELPTVVRNMGVGVSSVTSASRLGSILS 477
 DB 421 qlvpdlnhlystlvmvgkfgitaysmvvyvtaelptvvrnmvgvsvstasrlgsils 480
 QY 478 PFVYLGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDITDQMLRVKGMKHRTPS 537
 DB 481 pfvylgaydrflpyilmgsiltaitlffpessgvslpetidemqkvkklqrgsls 540
 RESULT 7
 AAY01649
 ID AAY01649 standard; Protein; 551 AA.
 XX
 XX AAY01649;
 XX
 XX 23-JUN-1999 (first entry)
 XX
 XX A protein with cation transporting activity.
 XX
 XX Organic cation transporter; OCT1; OCT2; drug development; fatty liver;
 KW heart disease; cancer; anti-tumour drug; anticancer drug.
 XX
 XX Homo sapiens.
 XX
 XX WO9913072-A1.
 XX
 XX 18-MAR-1999.
 XX
 XX 07-SEP-1998; 98WO-JP04009.
 XX
 XX 20-MAY-1998; 98JP-0156660.
 XX 08-SEP-1997; 97JP-0260972.
 XX
 XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
 XX
 XX Nezu J, Oku A;
 XX
 XX WPI: 1999-215062/18.
 XX N-PSDB; AAX26879.
 XX
 XX Genes homologous with organic cation transporters OCT1 and OCT2,
 PT useful in design of new drugs for treatment of diseases due to
 PT abnormality of the transporter functions
 XX
 XX Claim 1; Page 41-45; 97pp; Japanese.
 XX
 XX The present sequence represents a protein with cation transporting
 CC activity. The genes are significantly homologous with organic cation
 CC transporters OCT1 and OCT2. The genes may be used in drug development,
 CC particularly in the treatment of diseases due to abnormality of the
 CC organic cation transporter functions e.g. fatty liver, heart diseases
 CC and cancers, by controlling such as by inhibition or activation.
 CC Administration of anti-tumour and anticancer drugs in combination with
 CC a transporter protein inhibiting agent allows the agents to penetrate
 CC into the diseased cells to enhance the drug action.
 XX
 XX Sequence 551 AA;
 XX
 XX Query Match 76.9%; Score 2218; DB 20; Length 551;
 XX Best Local Similarity 75.9%; Pred. No. 4e-233;
 XX Matches 423; Conservative 57; Mismatches 71; Indels 6; Gaps 3;
 QY 1 MRDYDEVTAFLGEMGPFORLFFLLSIIIPNGFTGLSSVFLIATPEHRCRVPDAANLSS 60
 DB 1 mldydevtaflgwtqfqlifllsaiipngftglssvflaipaehrcrripdvnlss 60

Db 1 mrdydeviafgevgpqrllifllsasiipngfngmsvflagtpehrcrvpdaanlss 60
Qy 61 AWRNHTVPLRLRDGQREVPHSCRRYRLATIANFSALEPGSDVDLGOLEQESCLDGEHFS 120
Db 61 awrnsvplrlrdgrevphscrryrlatiansaigleprgdvdlgledescldgwehs 120
Qy 121 QDYLSTIVTEWNLVCEDDWKAPLTISLFFVGLLGSFISQSLDRFGKRNVLFTVMGMQ 180
Db 121 qdylstivtewnlvcednwkvplttlsffvgvllgsfvsqgsdrgrknvlfatmavq 180
Qy 181 TGFSFLOIFSKNFEMFVFLVGVGMGOISNYAAVAFVLGTEILGKSVRIIPSTLGVCIFYA 240
Db 181 tgfsfqlfksnfmfvflvvgmgqisnyvavfllgtellgksvriifstlgvctffa 240
Qy 241 FGYNVPLFAYFIRDWRMLVALTMPGVLCAVLMWFIPESPRWLISQGRPEAEVIIRKA 300
Db 241 vgymlplfayfirdwrmllaltvpylcvplwfwfipesprwlisqrfrfreaedliqka 300
Qy 301 AKANGIVVPSTIFDPSELODSSKKQOSHNLDDLRTWNIRMTIMSLMWTISVGYFG 360
Db 301 akmntavpavifd--sveelnplkqgkafildfrtrniatimslmllwmltsvgyfa 358
Qy 361 LSLDTPNLHGDIFFVNCFLSAMVEVPAYVLAWLLOVLPARYSMATALFLGGSVLLFMQLV 420
Db 359 lsdaephldgdaylncflsaliepavitaillrtpryiaaavlfwg99villfiqlv 418
Qy 421 PPDIYLIATVLMVGKFGVTAAFSMVYVYTAELYPTVVRNMGVGSSTASRLGSILSPYF 480
Db 419 pvdyyfislglvmlgkfgitsafsmlyvftaelypvltrnmavgtstasrvgsiiapyf 478
Qy 481 VYLGAIDRFPLPYILMGSLTILTAILFLPESFGTLPDPTIDQMLRVKGMKHKRTPSHTR 540
Db 479 vylgaynrmipyilmgsiltvligitffpesigmtlpetleqmkvkwfrsgk---ktr 535
Qy 541 MLKDGQERPTILKSTAF 557
Db 536 dsmeteenpkvl-itaif 551

RESULT 8
AAY01651
ID AAY01651 standard; Protein; 553 AA.
AC AAY01651;
XX
DT 23-JUN-1999 (first entry)
XX
DE A protein with cation transporting activity.
XX
KW Organic cation transporter; OCT1; OCT2; drug development; fatty liver;
KW heart disease; cancer; anti-tumour drug; anticancer drug.
XX
OS Mus musculus.
XX
PN W09913072-A1.
XX
PD 18-MAR-1999.
XX
PF 07-SEP-1998; 98WO-JP04009.
XX
PR 20-MAY-1998; 98JP-0156660.
PR 08-SEP-1997; 97JP-0260972.
XX
PA (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
PI Nezu J, Oku A;
XX
DR WPI; 1999-215062/18;
DR N-PSDB; AAX26898.
XX
PT Genes homologous with organic cation transporters OCT1 and OCT2,
PT useful in design of new drugs for treatment of diseases due to
PT abnormality of the transporter functions

XX
PS
XX
CC The present sequence represents a protein with cation transporting
CC activity. The genes are significantly homologous with organic cation
CC transporters OCT1 and OCT2. The genes may be used in drug development,
CC particularly in the treatment of diseases due to abnormality of the
CC organic cation transporter functions e.g. fatty liver, heart diseases
CC and cancers, by controlling such as by inhibition or activation.
CC Administration of anti-tumour and anticancer drugs in combination with
CC a transporter protein inhibiting agent allows the agents to penetrate
CC into the diseased cells to enhance the drug action.
XX
SQ Sequence 553 AA;

Query Match 74.5%; Score 2148; DB 20; Length 553;
Best Local Similarity 72.2%; Pred. No. 1.8e-225;
Matches 402; Conservative 67; Mismatches 84; Indels 4; Gaps 2;
Qy 1 MRDYDEVIAFGEWGPQRLLIFLLSASIIPNGFTGLSSVFLIATPEHRCRVDPDAANLSS 60
Db 1 mrdydeviafgevgpqrllifllsasiipngfngmsvflagtpehrcrvpdtvnls 60
Qy 61 AWRNHTVPLRLRDGQREVPHSCRRYRLATIANFSALEPGSDVDLGOLEQESCLDGEHFS 120
Db 61 awrnhsipletkdgqvpsqrryrlatiansaigleprgdvdlgledescldgweyd 120
Qy 121 QDYLSTIVTEWNLVCEDDWKAPLTISLFFVGLLGSFISQSLDRFGKRNVLFTVMGMQ 180
Db 121 kdiflsvtewnlvceddwktplttsifvlgsvsgslsdrgrkvvlfatmavq 180
Qy 181 TGFSFLOIFSKNFEMFVFLVGVGMGOISNYAAVAFVLGTEILGKSVRIIPSTLGVCIFYA 240
Db 181 tgfsfvqlfksnfmfvflvvgmgqisnyvavfllgtellgksvriifstlgvctffa 240
Qy 241 FGYNVPLFAYFIRDWRMLVALTMPGVLCAVLMWFIPESPRWLISQGRPEAEVIIRKA 300
Db 241 lgymlplfayfirdwrmllaltvpylcvplwfwfipesprwlisqrfrfreaedliqka 300
Qy 301 AKANGIVVPSTIFDPSELODSSKKQOSHNLDDLRTWNIRMTIMSLMWTISVGYFG 360
Db 301 akmsivapagifdplelqelskqkvildfrtrniatimslmllwmltsvgyfa 360
Qy 361 LSLDTPNLHGDIFFVNCFLSAMVEVPAYVLAWLLOVLPARYSMATALFLGGSVLLFMQLV 420
Db 361 lslnvplhgdvynclfsliievpayftaillrtpryiaaavlfwg99villfiqlv 420
Qy 421 PPDIYLIATVLMVGKFGVTAAFSMVYVYTAELYPTVVRNMGVGSSTASRLGSILSPYF 480
Db 421 pedyfvsvigvmlgkfgitsafsmlyvftaelypvltrnmavgtstasrvgsiiapyf 480
Qy 481 VYLGAIDRFPLPYILMGSLTILTAILFLPESFGTLPDPTIDQMLRVKGMKHKRTPSHTR 540
Db 481 vylgaynrmipyilmgsiltvligitffpesfgvltlpetleqmkvkrgfcgk---kst 537
Qy 541 MLKDGQERPTILKSTAF 557
Db 538 vsvdreespkvl-itaif 553

RESULT 9
AAB20579
ID AAB20579 standard; Protein; 553 AA.
XX
AC AAB20579;
XX
DT 11-DEC-2000 (first entry)
XX
DE Mouse OCTN1 amino acid sequence.
XX
KW Mouse; transporter; OCTN1; OCTN2; OCTN3; organic cation transporter;
KW identification; regulator; carnitine transport.

XX OS Mus musculus.
XX PN WO200046368-A1.
XX PD 10-AUG-2000.
XX PF 04-FEB-2000; 2000WO-JP00619.
XX PR 05-FEB-1999; 99JP-0028406.
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX PA Nezu J, Ose A;
XX PI WPI; 2000-586982/55.
XX DR
XX PT Organic cation transporter gene OCTN3 expressed in testis for
XX PT identification of regulators of carnitine transport for use as drugs
XX PS
XX PS Example 2; Fig 2; 58pp; Japanese.
XX The present invention describes a mouse organic cation transporter
XX protein (OCTN3). Also described are: (1) a method for screening
XX compounds for their ability to regulate the transport of an organic
XX cation into the cell, by generating a cell expressing OCTN3 at the
XX cell membrane, contacting with the compound and organic cation, and
XX observing the degree of transport of the organic cation; and (2) a
XX method for screening compounds for their ability to be transported into
XX the cell by OCTN3, by generating a cell expressing OCTN3 at the cell
XX membrane, contacting with the compound and observing the degree of
XX transport of the compound. OCTN3 can be used for the identification of
XX regulators of the transport of organic cations (especially carnitine)
XX into cells by OCTN3, for use as drugs. The present sequence represents
XX the mouse OCTN1 amino acid sequence, which is used in an example from
XX the present invention.
XX SQ Sequence 553 AA;

Query Match 74.5%; Score 2148; DB 21; Length 553;
Best Local Similarity 72.2%; Pred. No. 1.8e-225;
Matches 402; Conservative 67; Mismatches 84; Indels 4; Gaps 2;

QY 1 MRDYEVTALGEGWGPQRILIFLLSASIPNGFTGLSSVFLIATPEHRCRVPDANLSS 60
DB 1 mrdyeviaflegwgpqrilifllsasiipngfngmvsvflagtpehrcrlvdpdvnlss 60
QY 61 AWRNHTVPLRLRDGEPVPHSCRRYRLATIANFSGALGPEGRDVLGQLESCLDGWEFS 120
DB 61 swrnhstplektdgrvppqscrryrlatiansamflegpgdvldleqescldgweyd 120
QY 121 ODVYLSITVTENNLVCEDDNKAPITISLFVGVLLGSFISQQLSDRFRKNVLFTVMGQ 180
DB 121 kdifstivtennlvceddwktpittslffvlgcgsfvgglsdrfrkknvlfatmavq 180
QY 181 TGFSELFQFSKNPEFVFLVLCWGOISNVAAVFLGTTELCKSVRIIFSTGLGVCIFYA 240
DB 181 tgfsvlfqfscnpefvlvflcwggoisnvvaflgttelcksvrllfstlglvcffia 240
QY 241 FGYNVPLFAFYFIRDRLMLVALTPMGVLCVALWVFIPESPRWLISQGRFEAEVIRKA 300
DB 241 igymvplfayfirdwrmllaltlplfcvplwffipesprwlisqrfaaeaiqlka 300
QY 301 AKANGIVVPSTIFDPSELDLSSKKQSHNLLDLRTWNTIMVMTIMSLMWISVGYFG 360
DB 301 akmsivapagifdplelqelsikgkvilidfrtniatitvmavlmwmtsvgyfa 360
QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLALLQLYLPVRYSMATLFLGSLVLLFMQLV 420
DB 361 lslavpnlhgdvynclslgllievpayftawllrtpryilagvlfwgg99vllliqvv 420
QY 421 PPDLYLATVLMVGKFGVTAFAFAMVYVYTAELYPTVVRNMGVGSSTASRLGSILSPYF 480

DB 421 pedyfnfysiglvmlgkfgitsafsmlyvftaelyptlvrmnmgvitsmasrvgsilapvf 480
QY 481 VYLGAIDRFLPYILMGSILTILTAILTLFLPESFGTPLPDTIDQMLRVKGMKHKRTSHTR 540
DB 481 vylgaynrllipyilmgsltlvltliffbesfgvtlpenleqmqkvgfrcgk---kst 537
QY 541 MLKDGQERPTILKSTAF 557
DB 538 vsvdreespkvl-ita 553

RESULT 10
AAB43038
ID AAB43038 standard; Protein; 560 AA.
XX AC AAB43038;
XX DT 08-FEB-2001 (first entry)
XX DE Human ORFX ORF2802 polypeptide sequence SEQ ID NO:5604.
XX KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; erythematoidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.
XX OS Homo sapiens.
XX PN WO200058473-A2.
XX PD 05-OCT-2000.
XX PF 31-MAR-2000; 2000WO-US08621.
XX PR 31-MAR-1999; 99US-0127607.
XX PR 02-APR-1999; 99US-0127636.
XX PR 05-APR-1999; 99US-0127728.
XX PR 30-MAR-2000; 2000US-0540763.
XX PA (CURA-) CURAGEN CORP.
XX PI Shinkets RA, Leach M;
XX WI WPI; 2000-602362/57.
XX DR N-PSDB; AAC77247.
XX PT Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 4791-4792; 5507pp; English.
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
XX osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
XX antithyroid; and antianaemic. The sequences can be used for determining
XX the presence of or predisposition to, or preventing or treating

Db	241	gmksr-twasvnhshffavgtllvaltgylvrtwlygmilstvtvpfilcc-----wvlpe	296
QY	280	SPRWLSQGRFEAEVIRKAANKANGIVVPSTIFDPSELDL-----SSKKQOSHNI	331
Db	297	tpfwllsegryeaaqkivdimakwn-----rasscklsellsldlqgpvsnsptevqkhl	352
QY	332	LDLLRTNIRNMTIMLWMTISGVFGSLSDTPNHLGDFVNCFLSAMVEVPAYVLAW	391
Db	353	syfynwsitkrtltwvllwftslgfsfslnsnlgnneylnlllgvveipaytfvc	412
QY	392	LLQYLPRRYSMATALFLGGSVLLFQVLVPPDLYLYLATVLMVGK-FGYTAAFSMVYVT	450
Db	413	iamdkvgrtvlayslfcslacgvvmvpqkhyilgvvtamgkilpigaagilylt	472
QY	451	AELYPTVVRNMGVSVSTASRLGSLSPYFYVYLGAIDRFLPYILMGLSLTLTLLFLP	510
Db	473	aelyptivrsilavsgsmvcrslapfsvdlssiwifipqlfvgtmallsgvltklp	532
QY	511	ESFGTPLPDTIDQMLRVKGMKHKRT	535
Db	533	etlgkrlattweeaakiesenesks	557
RESULT 12			
Id	AAW64538	standard; Protein; 554 AA.	
XX	AAW64538;		
XX	21-OCT-1998	(first entry)	
XX	Human liver cell clone HP01293	protein.	
XX	Transmembrane domain; human; nutrition; cytokine; cell proliferation;		
XX	differentiation; immune system; stimulator; suppressor; regulator;		
XX	hematopoiesis; activin; inhibitor; chemotactic; chemokinetic; receptor;		
XX	haemostatic; thrombolytic; ligand; anti-inflammatory; tumour.		
XX	Homo sapiens.		
XX	WO9821328-A2.		
XX	22-MAY-1998.		
XX	07-NOV-1997;	97WO-JP04056.	
XX	13-NOV-1996;	96JP-0301429.	
XX	(PROT-) PROTEGENE INC.		
XX	(SAGA) SAGAMI CHEM RES CENTRE.		
XX	Kato S, Kobayashi M, Sekine S, Yamaguchi T;		
XX	WPI; 1998-297932/26.		
XX	N-PSDB; AAV49558, AAV49559.		
XX	Human protein having transmembrane domain - useful for, e.g.		
XX	research and nutrition		
XX	Claim 1; Page 96-98; 205pp; English.		
XX	AAW64534-W64558 represent human proteins containing a transmembrane		
XX	domain. These proteins can be used for, e.g. research and nutrition, and		
XX	may have cytokine and cell proliferation/differentiation, immune		
XX	stimulating/suppressing, haematopoiesis regulating, tissue growth,		
XX	activin/inhibin, chemotactic/chemokinetic, haemostatic and thrombolytic,		
XX	receptor/ligand, anti-inflammatory or tumour inhibition activity.		
XX	Sequence	554 AA;	

Query Match	25.5%;	Score 735;	DB 19;	Length 554;
Best Local Similarity 35.8%; Pred. No. 3.8e-71;				
Matches 196; Conservative 86; Mismatches 216; Indels 50; Gaps 12;				
QY	1	MRDYDEVTAFLGEGWPPFOR---LTFLLSASLIIPNGTGLSSVFLIATPEHRCRVPPDAAN	57	
Db	1	mptvddillegvsgwfqkqafllilclisaafap---icvgivflgftpdhhcdspgvae	57	
QY	58	LSS--AWR-----NHTVPLRLRDGREVPHSCRRYR-----LATIA-NFSALG	96	
Db	58	lsqrcgswpaeeelnvypglpageaflgqcrreyevdwngalscscvdpplasiatnrshp	117	
QY	97	LBPGRDVLGQLEQBSCLDGMWFSQDVLVSTIVTEWNLVCEDDMKAPLTISLFFVGVLLG	156	
Db	118	lqp-----cqdgwvy--dtpgssivtefnlvcadskldlfqscnagffg	162	
QY	157	SFISQSLSDRFGKKNVLFVTMGMTQGFSELFQIFSKNFEMFVFLVVGMGQISNYYAAV	216	
Db	163	slgvygfadrfrgkcllgtvlvnavsgvlmafspnysmlfrllqglvskgnmagyt	222	
QY	217	LGTEILGKSVRIESTLGVCIYAF--GYMVLPLFAYFIRDRWRLVLTWPGVLCVALM	274	
Db	223	litefvgsgr---rtvaimygmavtvgvalgtlayalphwrwlqlavslptflilly	279	
QY	275	WFIPESPRWLISQGRFEAEVIRKAANKANGIVVPSTIFDPSELDLSSKKQOSHNI	334	
Db	280	wcvpesprwllsqkrnteaikimdhiaqngklppadikmlsleedvtek--lspsfadl	337	
QY	335	LRTWIRNIRNMTIMLWMTISGVFGSLSDTPNHLGDFVNCFLSAMVEVPAYVLAWLL	394	
Db	338	frtprlrkrtilmylftdsvlyqgllhmgatsgnlyldfysalveipgafialiti	397	
QY	395	QYLPRYSMATALFLGGSVLLFQVLVPPDLYLYLATVLMVGKFGVTAAFSMVYVTAELY	454	
Db	398	drvgrlypmavnnlagaacvlmifispldhwnliimcvmrgmitaigmiclvnaely	457	
QY	455	PTVVRNMGVSVSTASRLGSLSPYFYVYLGAIDRFLPYILMGLSLTLTLLFLPESF	513	
Db	458	ptfvrnlgymvcsslcdiggitpfivfrrevwqalpilfavgliaagvtllilpetk	517	
QY	514	GTPLPDTI	521	
Db	518	gvalpetm	525	
RESULT 13				
Id	AAW49401	standard; Protein; 540 AA.		
XX	AAW49401;			
XX	02-MAR-2001	(first entry)		
XX	Murine organic anion transporter 6.			
XX	Murine; organic anion transporter 6; mOATP6; cancer; inflammation;			
XX	cardiovascular disease; central nervous system disorder; kidney disease;			
XX	liver disease; autoimmune disease.			
XX	Mus sp.			
XX	WO200070048-A1.			
XX	23-NOV-2000.			
XX	15-MAY-2000;	2000WO-US13316.		
XX	14-MAY-1999;	99US-0134137.		
XX	12-MAY-2000;	2000US-0570293.		
XX	(SMIK) SMITHKLINE BEECHAM CORP.			
XX	(SMIK) SMITHKLINE BEECHAM PLC.			
XX	Feild J, Yue L, Ellens H;			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:10:16 ; Search time 23.33 Seconds
(without alignments)
537.262 Million cell updates/sec

Title: US-09-521-195-3
Perfect score: 2883
Sequence: 1 MRDYDEVTAFLGNGPQRL.....HTRMLKDGQRPILKSTAF 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	726	25.2	555	US-08-501-572-3	Sequence 3, Appli
2	726	25.2	555	US-09-040-444-3	Sequence 3, Appli
3	718	24.9	556	US-08-501-572-1	Sequence 1, Appli
4	718	24.9	556	US-09-040-444-1	Sequence 1, Appli
5	714.5	24.8	553	US-08-501-572-2	Sequence 2, Appli
6	714.5	24.8	553	US-09-040-444-2	Sequence 2, Appli
7	652.5	22.6	537	US-08-647-397-2	Sequence 2, Appli
8	302.5	10.5	520	US-08-964-127-2	Sequence 2, Appli
9	234	8.1	494	US-09-031-392-5	Sequence 5, Appli
10	234	8.1	494	US-09-299-549-5	Sequence 4, Appli
11	214.5	7.4	286	US-08-964-127-4	Sequence 3, Appli
12	208	7.2	492	US-08-355-844-3	Sequence 3, Appli
13	208	7.2	492	PCT-US95-16126-3	Sequence 3, Appli
14	203	7.0	493	US-09-031-392-10	Sequence 10, Appl
15	203	7.0	493	US-09-299-549-10	Sequence 10, Appl
16	198	6.9	109	US-08-647-397-4	Sequence 4, Appli
17	193	6.7	509	US-09-031-392-6	Sequence 6, Appli
18	193	6.7	509	US-09-299-549-6	Sequence 6, Appli
19	186	6.5	500	US-09-031-392-7	Sequence 7, Appli
20	186	6.5	500	US-09-299-549-7	Sequence 7, Appli
21	184	6.4	563	US-09-031-392-2	Sequence 2, Appli
22	184	6.4	563	US-09-299-549-2	Sequence 2, Appli
23	175	6.1	488	US-08-928-692-11	Sequence 11, Appl
24	171	5.9	322	US-08-964-127-6	Sequence 6, Appli
25	168.5	5.8	534	US-09-031-392-4	Sequence 4, Appli
26	168.5	5.8	534	US-09-299-549-4	Sequence 4, Appli
27	155	5.4	383	US-09-031-392-3	Sequence 3, Appli

28	155	5.4	383	4	US-09-299-549-3	Sequence 3, Appli
29	153.5	5.3	524	2	US-08-928-692-12	Sequence 12, Appl
30	149	5.2	584	2	US-08-928-692-13	Sequence 13, Appl
31	131	4.5	488	2	US-08-928-692-10	Sequence 10, Appl
32	110	3.8	473	1	US-08-597-236-13	Sequence 13, Appl
33	110	3.8	473	1	US-08-746-682A-13	Sequence 13, Appl
34	104.5	3.6	1835	3	US-08-836-325-15	Sequence 15, Appl
35	101.5	3.5	341	1	US-08-423-564-5	Sequence 5, Appli
36	97	3.4	342	3	US-08-785-928-1	Sequence 1, Appli
37	97	3.4	342	3	US-08-728-603-17	Sequence 17, Appl
38	97	3.4	916	2	US-08-928-692-58	Sequence 58, Appl
39	94.5	3.3	390	5	PCT-US92-02091-8	Sequence 8, Appli
40	93.5	3.2	467	2	US-08-805-118-3	Sequence 3, Appli
41	93.5	3.2	480	2	US-08-724-394A-9	Sequence 9, Appli
42	93	3.2	560	2	US-08-805-118-4	Sequence 4, Appli
43	92.5	3.2	419	1	US-08-385-186-2	Sequence 2, Appli
44	92.5	3.2	568	4	US-08-637-823B-27	Sequence 27, Appl
45	92.5	3.2	584	4	US-08-637-823B-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-08-501-572-3
; Sequence 3, Application US/08501572
; Patent No. 6063623
; GENERAL INFORMATION:
; APPLICANT: Koepsell, Hermann
; APPLICANT: Grundeman, Dirk
; TITLE OF INVENTION: Transport protein Which Effects The
; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,
; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson,Parabow,Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/501,572
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Toohy, Kimberlin M
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 02481.1453-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 555 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-501-572-3

Query Match 25.2%; Score 726; DB 3; Length 555;
Best Local Similarity 35.0%; Pred. No. 5.6e-68;
Matches 187; Conservative 98; Mismatches 223; Indels 26; Gaps 14;

Db	6	DDVLEHGGEFHFKQKQFFLLALLSATFAPY-YVGI--VELGFTPHRCRSPGVAELSLR	62
Qy	60	SAWR-----NHTVPLRLRDGVEPHCHRRYRLA-TIANFSALGLEPGROVDLQOLQESQ	113
Db	63	CGWSPAEBELNYTVGPGPAGEASPRQCRREYVDWNOSTFCDVDPDLASLOTNRSLRPGPC	122
Qy	114	LDGNEFSODVYLSTIVTEWNLVCEDDKKAPLTTISLEFVGVLGSGFISGQLSDRFRGNVL	173
Db	123	RDGWY--ETPGSSIVTEFNLCANSWMLDFQSSNVNGFFIGSMSTGYIADREGRKCL	180
Qy	174	FVTMGMTGSFLQIFSKNPFEMFVLEVLVGMQGISNYVAAFVLTGTEILGKSVRIITFSL	233
Db	181	LTTVLIINAAGVLMASPTVTWMLIFRLIOGLVSKAGWLIGYTLITFEFVGGRYR--RTV	237
Qy	234	GVC--IFYAFGYMPLPFAFYIDWRMLLVALTMPGVLCAVMWFPIPESPRMLISOGREF	291
Db	238	GIFYQAVTYVGLLVLAGVAVALPHRWLQFTVALPNFFFLLYKWCIPESPRMLISONKNA	297
Qy	292	EAELVIIRAKAKANGIVVPSTIFDPSELQDLSKKQKSHNILDLLRTWNTIRMTXIMLW	351
Db	298	EAMRIIHKIAKKNCKSLPASL-ORLLEETGKK-LNPSFLDLVRTQPKRKHMTMLMYNW	355
Qy	352	MTISGVGFGLSDTPNLHGD-IFVNCFLSAMVEVPAYVLAWLLQLVLPYRYSWATALFUG	410
Db	356	FTSSVLQGLIMHN-GLAGONIYLDFFYSALVEFPAAFMILIIDRGRYRYPWAAASNMYA	414
Qy	411	GSVLLFMQLVPPDLYLATVLVMVGKFGVTAAPSMYVYVTAELYPTVWNMGVGSYSTAS	470
Db	415	GAACLASVFI PGDLOWLKIISCLGRMGITMAYEIVCLVNAELYPTFINLGVHVCSSMC	474
Qy	471	RLGSLSPFYFY-IGAYDRELPIYLMGSLTILTAITLFLPSFGSGPLPDTIDQ	523
Db	475	DIGIIPFYRLTNLWELPMFVGLVAGGLIALLPTKGAALPETEE	528

RESULT

US-09-040-444--3
; Sequence 3, Application US/09040444
; Patent No. 6063766
; GENERAL INFORMATION:
; APPLICANT: Koepsell, Hermann
; APPLICANT: Grundeman, Dirk
; APPLICANT: Gorboulev, Valentin
; TITLE OF INVENTION: Transport protein Which Effects The
; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,
; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan,Henderson,Farabow,Garrett & Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,444
; FILING DATE: March 18, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Connor, Steven P
; REGISTRATION NUMBER: 41,225
; REFERENCE/DOCKET NUMBER: 2481.1453-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/501,572
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Toohay, Kimberlin M
; REGISTRATION NUMBER: 35,391
; REFERENCE/DOCKET NUMBER: 02481.1453-00000
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-501-572-1

```

```

Query Match 24.9%; Score 718; DB 3; Length 556;
Best Local Similarity 34.0%; Pred. No. 3,9e-67;
Matches 194; Conservative 96; Mismatches 229; Indels 52; Gaps 14;

```

```

QY 1 MRDYDEVTAFLGEGWGFQRLIFLL---SASIIPIGFTGLSSVFLIATPEHRCRVPDAA 57
Db 1 MPTVDDVLEQGEFGFQKQAFLLCLISASLAPI-YVGI--VFLGFTPGHYCQNPV 57
QY 58 LSS--AWR-----NHTVP-LRLRDGREVPHSCRRYL-----ATIANFSAL 95
Db 58 LSQRCGWSQAEELNYTVPGLPDSDEASFLSQCMRYEVDWQSTLDCVDP LSSLVANRSQL 117
QY 96 GLEPGRDVDLQLEQESCLDGEFSDQVYLSITVETNVLCEDDWKAPLTISLFFVGVLL 155
Db 118 PLGP-----CEHGWY--DTPGSSIVTEFNLCVGDAMKVDLFOQCNVNGFLL 162
QY 156 GSFTSGQLSDRFRGKRVFLVTMGMOGTFSLQIFSKNFEMFVFLVVGNGOISNVA 215
Db 163 GSVVGYIADRFGRKCLLVTLTTSVSGVLTAVAPDYTSMLLFRLLQGVKSGVWSGY 222
QY 216 VLGTIELGKSVRIIFSTLGVCIYAF--GYMVLPLFAYFIRDMRLMLVALTMPGVL 273
Db 223 TLIEFVSGYR---RTAILYQMAFTVGLVGLAGVAYIPDWRWLQVLSLPTFLFLY 279
QY 274 WFTIPESRWLISQGRFEAEVIRKAAKANGIVVPTIFDPSLQDLSSKKQOOSHNL 333
Db 280 YWFPESRWLLSQKRTTRAVRIMEQIAQKNGKVPADLKLMLCLEEDASEKR--SPSFAD 337
QY 334 LLRTWNRIMVTIMSLMWMITISVGYFGLSLDTPNLHGDI FVNCFLSAMVEVPAYVLA 393
Db 338 LFRTPNLRKHTVILMYLWFSQAVLYQGLIMHVAGATGANLYLDFFYSLSVEFPAAFI 397
QY 394 LQYLPRIYSMATALFLGGSVLLFMQVPPDLYLYLATVLMVVGKFGVTAAFSMVYV 453
Db 398 IDRIGRIYPIAASNLVTGAACLLMIFIPHELHNLVTLACLGRMGATVLEWVCLVNA 457
QY 454 YPTVRRNMGVSGVSTASRLGSLSPFYV--LCAYDRFLPYLIMGSLILITLFLPES 512
Db 458 YPTFTRNLGMVCSALCDLGIFTFPMVFRMLMEVMOALPLILFGLVLTAGATLLLPET 517
QY 513 FCTPLPTIDQMLRVKGMKHKRTKPSHTRLK 543
Db 518 KGVALPTEIEAENL-GRKSKAKENTYILQ 547

```

```

RESULT 4
US-09-040-444-1
; Sequence 1, Application US/09040444
; Patent No. 6063766
; GENERAL INFORMATION:
; APPLICANT: Koepsell, Hermann
; APPLICANT: Grundeman, Dirk

```

```

; APPLICANT: Gorboulev, Valentin
; TITLE OF INVENTION: Transport protein which effects the
; TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals,
; TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,444
; FILING DATE: March 18, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Connor, Steven P
; REGISTRATION NUMBER: 41,225
; REFERENCE/DOCKET NUMBER: 2481.1453-01
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-040-444-1

```

```

Query Match 24.9%; Score 718; DB 3; Length 556;
Best Local Similarity 34.0%; Pred. No. 3,9e-67;
Matches 194; Conservative 96; Mismatches 229; Indels 52; Gaps 14;

```

```

QY 1 MRDYDEVTAFLGEGWGFQRLIFLL---SASIIPIGFTGLSSVFLIATPEHRCRVPDAA 57
Db 1 MPTVDDVLEQGEFGFQKQAFLLCLISASLAPI-YVGI--VFLGFTPGHYCQNPV 57
QY 58 LSS--AWR-----NHTVP-LRLRDGREVPHSCRRYL-----ATIANFSAL 95
Db 58 LSQRCGWSQAEELNYTVPGLPDSDEASFLSQCMRYEVDWQSTLDCVDP LSSLVANRSQL 117
QY 96 GLEPGRDVDLQLEQESCLDGEFSDQVYLSITVETNVLCEDDWKAPLTISLFFVGVLL 155
Db 118 PLGP-----CEHGWY--DTPGSSIVTEFNLCVGDAMKVDLFOQCNVNGFLL 162
QY 156 GSFTSGQLSDRFRGKRVFLVTMGMOGTFSLQIFSKNFEMFVFLVVGNGOISNVA 215
Db 163 GSVVGYIADRFGRKCLLVTLTTSVSGVLTAVAPDYTSMLLFRLLQGVKSGVWSGY 222
QY 216 VLGTIELGKSVRIIFSTLGVCIYAF--GYMVLPLFAYFIRDMRLMLVALTMPGVL 273
Db 223 TLIEFVSGYR---RTAILYQMAFTVGLVGLAGVAYIPDWRWLQVLSLPTFLFLY 279
QY 274 WFTIPESRWLISQGRFEAEVIRKAAKANGIVVPTIFDPSLQDLSSKKQOOSHNL 333
Db 280 YWFPESRWLLSQKRTTRAVRIMEQIAQKNGKVPADLKLMLCLEEDASEKR--SPSFAD 337
QY 334 LLRTWNRIMVTIMSLMWMITISVGYFGLSLDTPNLHGDI FVNCFLSAMVEVPAYVLA 393
Db 338 LFRTPNLRKHTVILMYLWFSQAVLYQGLIMHVAGATGANLYLDFFYSLSVEFPAAFI 397
QY 394 LQYLPRIYSMATALFLGGSVLLFMQVPPDLYLYLATVLMVVGKFGVTAAFSMVYV 453
Db 398 IDRIGRIYPIAASNLVTGAACLLMIFIPHELHNLVTLACLGRMGATVLEWVCLVNA 457

```



```
Best Local Similarity 35.4%; Pred. No. 9,1e-67;
Matches 194; Conservative 87; Mismatches 216; Indels 51; Gaps 13;

QY 1 MRDYDEVTAFLGCGPFR---LIFFLLSASLIIPNGFTGLSSVFLIATPEHRCRVPDPAAN 57
Db 1 MPTVDILEQVGSWFQKQAFILCLLSAFAF---ICVGIVFGLTTPDHHCOSPGVAE 57
QY 58 LSS---AWR-----NHTVPLRLRDGREVPHSCRRY-----LATIA-NFSALG 96
Db 58 LSQCGWSPAEELNVTVPGLCPAGEAFGLGCRRYEVDWNSALSQVDPPLASLATNRSLP 117
QY 97 LEPGRVDLGOLEQSCDUNEFSDVYLVSTVTEWNLVCEDDWKAPLTISLFFVGVLLG 156
Db 118 LGP-----CQDQWVY--DTPGSSIVTEFNLVCAWSKLDLFOCLNAGFFG 162
QY 157 SPISGOLSDRGRKNVLEVTMGOTGFSFLOIFESKNPEMFVFLVYMGQISNYAAV 216
Db 163 SLVGIFADRFGRKLCUGTLVNAVSVLMFSPNYSMILLFLOGLVSKGNWAGYT 222
QY 217 LGTEILGKSVRIIFSTLGVCFYAF--GYMVLPLFAFIRDRMLLVALTMPGVLCVALW 274
Db 223 LITEFVGSSR---RTVAIMQMAFTVGLVALTGLAYALPHRWLQLAVSLPTFLFLY 279
QY 275 WPIPSRWLISQGRFEEAEVIRKAAKANGIVVPSTIEDPSELQDLSKKQKQSHNILD 334
Db 280 MCVPESRWLISQKRNTEAIIKMDHIAQKNGKLPADLMLSLSEEDVTEK--LSPSEADL 337
QY 335 LRTWIRWMTIMSINLWMTISVGYEGLSLDTPNLHGDIFVNCFLSAMVEPAYVLAWLL 394
Db 338 FTPLRLKRTFILMYLWFTSDVLYOGLLHMGATSGNLYDLFLYSALVEIPGAFIALITI 397
QY 395 QYLPYRYSMATLFLGGSVLLFMQLVPPDLYLATVLMVVGKFGVTAAFSMVYVYTAELY 454
Db 398 DRVGRIYPMAYNSLLAGAACLVFI-FISDLHLNIIIMCVGEMGITAIQMICLVNAELY 456
QY 455 PTVWRMGVGSSTASRLGSLSPYFY--LGAYDRFLPYILMGSLTILTAITLFLPESF 513
Db 457 PTFVRLNRMVCGSSICDGGITPFIIVRFLRVWQALPLILFAVLGLIAAGVTLILPETK 516
QY 514 GTPLPDTI 521
Db 517 GDALPETM 524

RESULT 7
US-08-647-397-2
; Sequence 2, Application US/08647397
; Patent No. 5972702
; GENERAL INFORMATION:
; APPLICANT: Beier, David R.
; APPLICANT: Brady, Kevin P.
; TITLE OF INVENTION: OSTEOCLAST TRANSPORTER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/647,397
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Gates, Edward R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: B0801/7048

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 537 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-647-397-2

Query Match 22.6%; Score 652.5; DB 2; Length 537;
Best Local Similarity 29.9%; Pred. No. 3e-60;
Matches 167; Conservative 104; Mismatches 246; Indels 41; Gaps 12;

QY 4 YDEVTAFLGEGWGFQRLIFLLSASLIIPNGFTGLSSVFLIATPEHRCRVPDPAANLSSAWR 63
Db 3 FSEILDRVSGMGPFQYLHVTLALPILGIANHNLLQIFTATTDPDHCRRPPNASL-EPW- 60
QY 64 NHTVPLRLRDGREVPHSCRRY---RLATIANFSALGLEPGRDVLGQLEQSCLDGWEEFS 120
Db 61 --VLPLG-PNGK--PEKCLRFVHLPHNASLPNDTQGATEP-----CLDGIWYN 102
QY 121 QDVYLSITVTEWNLVCEDDWKAPLTISLFFVGVLLGSFISGQSLDRFRKNVLFVTMGMQ 180
Db 103 ST--RDTIVTEWDLVCGSNKLEMAQSVFMAGILVGPVFGELSDRGRKPILTWSYLLL 160
QY 181 TGRSFLOIFESKNPEMFVFLVYMGQISNYAAVFLGTLEILGKSVRIIFS-TLGVCIIFY 239
Db 161 AASGSSAAAFSPSLTVYIMIFRFLGCGSISGISLSTIILNVEWVPTSTRAISSTTIGYC--Y 218
QY 240 AFGYMWLPFAFIRDRMLLVALTMPGVLCVALWFIPESPRWLISQGRFEEAEVIRK 299
Db 219 TICQFPLGLAYAYVQWRWQLSVSAAFIFSLLSWVPEIRWLVLSKFSRALKTLOR 278
QY 300 AAKANGIWEPTIEDPSEL-----QDLSKKQKQSHNILDRLTWNIRWMTIMSINLWMTI 354
Db 279 VATFNGKKEEGEKLTVELKFNQLQKITSKVK-YGLSDLFVRSILRRVTFCLSLAWFAT 337
QY 355 SVGYEGLSLDTPNLHGDIFVNCFLSAMVEPAYVLAWLLLOYLPRRYSMATLFLGGSVL 414
Db 338 GFAYYSIANGWVEEFGVNIYILQIFGGVDIPAKFITILSISYLGRRITQGGFLILLAGVAI 397
QY 415 LFMQLVPPDLYLATVLMVVGKFGVTAAFSMVYVYTAELYPTVVRNMGVGSSTASRLGS 474
Db 398 LALIFVSSEMQLLRTALAVEGKGLCSGFSCLFYSLEYPTVLRGTMGISINIRAVGS 457
QY 475 ILSPYFVYLGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDITID-----QMLRV 527
Db 458 MIAPLVKITGELQFPFIPNVIPTWMTLIGSSAAFFLLETNLRPLPETIEDIQDWYQQTKT 517
QY 528 KGMKHKRTPSHTRMLKDG 545
Db 518 KQEPFAEKASQTIPLTKTG 535

RESULT 8
US-08-964-127-2
; Sequence 2, Application US/08964127
; Patent No. 6277565
; GENERAL INFORMATION:
; APPLICANT: Grandearl, Andrew David John
; TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
```

```

, , COMPUTER READABLE FORM:
, , MEDIUM TYPE: Diskette
, , COMPUTER: IBM Compatible
, , OPERATING SYSTEM: Windows 95
, , SOFTWARE: FASTSQ for Windows Version 2.0b
, , CURRENT APPLICATION DATA: US/08/964,127
, , APPLICATION NUMBER: US/08/964,127
, , FILING DATE: 06/NOV/1997
, , PRIOR APPLICATION DATA:
, , APPLICATION NUMBER:
, , FILING DATE:
, , ATTORNEY/AGENT INFORMATION:
, , NAME: Crews, Ph.D., L. Lee
, , REGISTRATION NUMBER: P-43,567
, , REFERENCE/DOCKET NUMBER: 07334/038001
, , TELECOMMUNICATION INFORMATION:
, , TELEPHONE: 617/542-5070
, , TELEFAX: 617/541-8906
, , TELEX: 200154
, , INFORMATION FOR SEQ ID NO: 2:
, , SEQUENCE CHARACTERISTICS:
, , LENGTH: 520 amino acids
, , TYPE: amino acid
, , TOPOLOGY: linear
, , MOLECULE TYPE: protein
, , FRAGMENT TYPE: internal
, ,
, , US-08-964-127-2

```

Query Match	10.5%;	Score 302.5;	DB 4;	Length 520;
Best Local Similarity	22.5%;	Pred. No. 2.5e-23;		
Matches 125;	Conservative	98;	Mismatches 231;	Indels 101;
				Gaps 20;

Qy	40	VFLIATPEHRCRVPD	DAANI	SSAWRNITVPLRL	DRG	REVP	PHSCRRY	KLATIANF	SALGLEP	99																																											
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																																												
Db	6	IFTLAPPLH	-CHYGAF	PPNAGS	QWEP	PNASG	VS	ASAALA	ASARVAT	SDPSCGFAP	64																																										
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																																												
Qy	100	GRVDLQ	LBQES	CLDGE	ESQDV	YLST	-IVT	ENWLV	CEDD	KAPLIT	ISLFFV	GLLSF	158																																								
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																																									
Db	65	P	-----	-DFN	CL	KD	WYNG	LP	UVT	NAIG	OWDLV	CDLGM	OVILEQ	TLF	ILGF	ASGYL	116																																				
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																																					
Qy	159	ISQ	SLDR	FRG	KNV	LV	TWGM	-----	-QT	GS	FLQ	IF	SKN	FMF	VV	LV	LVGM	205																																			
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																																					
Db	117	FLG	VP	ADR	FR	GR	RG	IV	LLT	LG	LV	PCG	VG	NA	AGS	ST	GV	VALR	-----	FL	LG	ELL	AGV	168																													
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																													
Qy	206	GO	IS	NV	Y	AA	V	FL	TE	IL	G	K	S	V	R	I	F	T	L	G	V	C	I	F	A	F	Y	I	R	D	R	M	L	I	V	A	L	T	M	265													
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :																
Db	169	D	-----	-L	G	V	L	M	R	E	L	C	D	T	Q	R	L	R	A	L	A	G	L	V	G	G	H	F	L	G	L	-A	L	V	S	K	D	R	F	L	O	R	M	I	T	A	222						
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :													
Qy	266	P	G	V	L	C	A	L	A	L	M	-	F	I	P	E	S	P	R	L	I	S	O	G	R	E	A	E	A	V	I	R	K	A	A	K	A	N	-----	G	I	V	P	S	T	I	F	D	S	E	L	O	320
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :			
Db	223	P	C	I	L	F	Y	G	M	P	G	L	F	E	S	A	R	L	I	V	K	R	O	E	A	S	V	L	R	I	A	E	R	N	P	H	O	M	L	G	E	E	A	-----	Q	E	A	L	O	D	280		
		: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :	: : : : :			
Qy	321	L	S	S	K	K	O	Q	S	H	N	I	L	D	L	E	T	W	N	I	R	M	V	T	I																												

RESULT 9
 US-09-031-392-5
 ; Sequence 5, Application US/09031392
 ; Patent No. 5942398
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Weng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
 ; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: Fastseq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/031,392
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 494 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-031-392-5

[illegible]

Db 395 QGPRPAAVAGCSNWTNLFVGLLPSATFYLGA1-----VFIVTFVLFVIFVWFTEFKV 450
QY 510 PSECTPLPDIDQMLRV-----KGMKHKRTP 536
Db 451 PETRGR---TFEITRAFEGVQTGTGRGKGP 479

RESULT 10
US-09-299-549-5
; Sequence 5, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,549
; FILING DATE: 26-APR-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 494 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-299-549-5

Query Match 8.1%; Score 234; DB 4; Length 494;
Best Local Similarity 24.7%; Pred. No. 3.9e-16;
Matches 112; Conservative 85; Mismatches 158; Indels 98; Gaps 23;

QY 144 LTISLFFVGVLLGSPISQSLDRGK-----NVLFVTMGMTQGFSLQIFSKNFEMFV 197
Db 65 LVSALFSGVGGIGSSVGLFNRGRNSMLVNLAIATAGGLMGFKCI---AESVEMLI 121

QY 198 VLFVLVGM--GQISNVAFAVLGTGILKSVRIIFSTL-----GVCIFVAFGYMVLPL 248
Db 122 LGRLLIGFCGLCTGFVPMYI--GEISTALRGAFGLNLQIGVIGILVAQIFGLKVI-- 177

QY 249 FAYFIRD--WRMLVALTPGVLCAVMWFIPESRW--LISQRFEEAEVIRKAAKANGI 306
Db 178 --LGTELDWLLPLLGITLPAIIQCAALPCFSPRELLINKKEEKAKELIQLRWGTEDV 235

QY 307 VVPSTIFDPSELDSSKKQOSH--NILDLLRTWNIRVMTTMSIMLWMT-----ISVG 358
Db 236 AQ-----DIQEMKDESRMSQEKQVTVLELFRAPNYRPIIISIMLQSLQSLSINAVFY 290

QY 359 FGLSLDTNHLGDFVNCFLSAMVEVPAY-----VLAWLLQLVLPRIYSNAT 405

Db 291 YSNGI-----FKDAGVQEPVATIGAGVVNTIFTVVSVELVERAGRR--TLH 335
QY 406 ALPLGG-----SVLLFMQLVPPDLY-----YLATVLVMVGKEGVTAAAFSMVTVYTAELYP 455
Db 336 LIGLGMAFCSILMTISLLLKONYSWMSFCIGAILVFAFFEIGPG-PIPWFIVAELFG 394
QY 456 TVVRNMGVGVSS---TASRLGSILSPYFV-YLGAYDRELPLPYILMGSILTILAILTLF-L 509
Db 395 QGPRPAAVAGCSNWTNLFVGLLPSATFYLGA1-----VFIVTFVLFVIFVWFTEFKV 450
QY 510 PSECTPLPDIDQMLRV-----KGMKHKRTP 536
Db 451 PETRGR---TFEITRAFEGVQTGTGRGKGP 479

RESULT 11
US-08-964-127-4
; Sequence 4, Application US/08964127
; Patent No. 6277565
; GENERAL INFORMATION:
; APPLICANT: Grandearl, Andrew David John
; TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,127
; FILING DATE: 06-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crews, Ph.D., L. Lee
; REGISTRATION NUMBER: P-43,567
; REFERENCE/DOCKET NUMBER: 07334/038001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-964-127-4

Query Match 7.4%; Score 214.5; DB 4; Length 286;
Best Local Similarity 24.6%; Pred. No. 1.9e-14;
Matches 69; Conservative 51; Mismatches 122; Indels 39; Gaps 8;

QY 40 VFLIATPEHRCRYPDRAANLSSAMRNHTVPLRLRDGREVPHSCRRYRLATIANFSALEP 99
Db 6 IFTIATPLH--CHYGATPPNAGSGWEPNAGSVGVSAAALAAASRAVATSTDFSCSGFAP 64

QY 100 GRDVLGQLQEQESCLDQWESFQDVYLST-IVTEWNLVCEDDMKAPLITISLFFVGVLLGSF 158
Db 65 P-----DFNCLDKWDYNGLPLVTNAIGQDWLDVCDLGNQVILEILFILGFASGYL 116


```

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meikiejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07304/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; UN-09-031-392-10

Query Match 7.0%; Score 203; DB 2; Length 493;
Best Local Similarity 25.0%; Pred. No. 7.3e-13;
Matches 111; Conservative 76; Mismatches 147; Indels 110; Gaps

QY 144 LTSLTFVGVLLGSFGISGQLSDRFGRK-----NVLEVTMGWOTGFSGFIQIFSKNFEVF 197
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 68 LSVSIFAIVGGMIGSFVLGXIGNLGRKXAMLVNNVLAIAAGLLMG---LAKXAASFEMLI 124
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 198 VLFVLVGM-----GOISNVAAFLVGLGTILGKSVRIIFFSTLGVCIFYAFGYM 244
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 125 LGRFIIGLYCGLSSGVPMYVGELSPALRGALGT-----LNQLGIVIGILIA-Q 173
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 245 VLPLFAYFTIR--WRMLLVALTMPGVLCAVMWFIPESPRWLI----SQGRPEEAEEVIIR 298
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 174 VLGLDLSLGNESLWPLLGLTGPVALLOLLLPFCPSRYLLINKNEARAKKALLQR 238
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 299 KAANGAVVPSTIFTDFSELQD---LSSKKQOSHNLDLRTWNTRWTVIMSIMLWMT- 353
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 234 GTADVSEV-----AEMKDSRXMKSEKXVS--VLEFRSRXYRQPVIIAIVQLSQ 283
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 354 ----ISGVYFGLSDTPNLHGDIFVNCFLSAMVEVPAY-----VLAWLLIQ 395
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 284 QLSGINAVEFYTSI-----FEKAGVGQPVYATIGAGVVNTVTVSVFVVE 330
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 396 YLPRYSMATALFLGG-----SVLLFMOLVPPDLYLATVLVMVKGCVTAAFS-----MV 446
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 331 RAGR--TLHLGLGGMACCAVLMTIALLDQVPWMXYSVISVAIFGFVAFFEVGPGPIP 388
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 447 VYVTAELYPTVRRNMNGVGS----TASRLGISLSPYFV-YLGAYDRFLPYIILMSITIL 501
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 389 WFIVAELFESQGPRAAATAVACFSNWTSNFIVGLLFOYIAELLGPY----VFIVEAVLLLL 444
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
QY 502 TAILT-LFLPESFGTPLDDTDQM 524
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db 445 FFIFTFLKVPETKGR---TFDEI 464
|::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||:

RESULT 15
UN-09-299-549-10
; Sequence 10, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10

```

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FastSeq for Windows Ver
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/299,549
 FILING DATE: 26-APR-1999
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/031,392
 FILING DATE: 26-FEB-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Meiklejohn, Ph.D., Anita L.
 REGISTRATION NUMBER: 35,283
 REFERENCE/DOCKET NUMBER: 07334/07
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 493 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-299-549-10

Query Match	7.08;	Score 203;	DB 4;	Length 493;
Best Local Similarity	25.08;	Pred. No. 7.3e-13;		
Matches 111;	Conservative	76;	Mismatches 147;	Indels 110; Gaps 21;
QY 144	LTISLFFGVGLGFSISQSLDRGRK-----NVLFVTMGMTQGFSLQIFSKNFEMFV	197		
Db 68	LSVSIFAVGGMIGSFLVXGIGNRLGRKXAMLVNVLAIAGLLMG-----LAXXAXSFEMLI	124		
QY 198	VLFLVGM-----GQISNVYAAFLVGLTGKSVRIFTSLGVCIIYAFAGYM	244		
Db 125	LGRITGLYCOLSSGVVPMYVGEISPTALRGALT-----LNOLGTVIGILIA-Q	173		
QY 245	VLPLFAYFIRD--WRMLLVALTMFGLVCVALWMTIPSPRWLI-----SOGFEBAEVIIR	298		
Db 174	VLGLDSLLGNESLWPLLLGLTGVPALLQLLLPCPESPRYLLINKKEARAKALQRLR	233		
QY 299	KAAKANGIVVPSTIPDSELOD-----LSSKKQOSHNLDLLRTWNIRMTVIMSIMLWMT-	353		
Db 234	GTADVSOEV-----AEMKDESRXXMKSEKXVS--VLELFRSRXYRQPVIIAIVLQSLQ	283		
QY 354	-----ISGVFGLSLDTPNLHGDFIVNCFLSAMVEVPAY-----VLAWLLLQ	395		
Db 284	QLSGINAVFYYSI-----PEKAGVGQPVYATIGAGVNTVFTVVSVFVWE	330		
QY 396	YLPFRYSNATALPLGG-----SVLLFMOLVPPDLYLATVLVMVGKFGVTAAFS-----MV	446		
Db 331	RAGRR--TLHLLGLGGMAGCAVLMTALALLDQVPMWSYVSIIVAFGVFAFFVGPGPIP	388		
QY 447	YVYTAELYPTVWRNMGVGS-----TASRLGSILSPYF-YLGAYDRFLPYILMGSLTIL	501		
Db 389	WFIVAEFLPSGPPRAIAVAGFSNWTSNFVGLGLFQVIAELGPGY-----VFIVFAVLLLL	444		
QY 502	TAIL-LFLPESFGTLPDPIDQM	524		
Db 445	FFITFLKVPETKGF-----TFDEI	464		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:09:47 ; Search time 28.31 Seconds
(without alignments)
1498.736 Million cell updates/sec

Title: US-09-521-195-3
Perfect score: 2883
Sequence: 1 MRDYDEVTAFLGEMGFQRL.....HTRMLKDGQERPTILKSTAF 557

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2883	100.0	557	JW0089	organic cation tra
2	2496	86.6	557	JE0346	high-affinity carn
3	777	27.0	576	T22509	hypothetical prote
4	733	25.4	593	JC4884	organic cation tra
5	712	24.7	556	S50862	organic cation tra
6	589	20.4	794	T27870	hypothetical prote
7	547.5	19.0	745	T16565	hypothetical prote
8	522	18.1	527	T01019	transport protein
9	483	16.8	539	C96758	probable protein
10	470.5	16.3	518	B86299	hypothetical prote
11	411	14.3	528	T21682	hypothetical prote
12	382.5	13.3	540	T25851	hypothetical prote
13	380.5	13.2	515	B96825	hypothetical prote
14	368.5	12.8	521	H86298	hypothetical prote
15	355.5	12.3	529	T23190	hypothetical prote
16	341.5	11.8	454	F75380	probable sugar tra
17	316	11.0	455	B83213	probable MFS trans
18	308	10.7	469	G65058	hypothetical prote
19	298.5	10.4	724	I50531	transmembrane tran
20	295.5	10.2	423	S74046	probable sugar tra
21	291.5	10.1	435	T15290	hypothetical prote
22	287	10.0	400	C69757	transporter homolo
23	287	10.0	422	G72234	hypothetical prote
24	279	9.7	524	T27082	hypothetical prote
25	274	9.5	683	S34961	synaptic vesicle p
26	272	9.4	478	T33985	hypothetical prote
27	271	9.4	480	T23608	hypothetical prote
28	267	9.3	443	E64725	yaaU protein - Esc
29	267	9.3	443	H85485	probable transport

30	266	9.2	423	2	T19030	hypothetical prote
31	265	9.2	459	2	G64937	hypothetical prote
32	265	9.2	459	2	H85787	probable transport
33	262	9.1	442	2	A83122	probable MFS trans
34	261.5	9.1	742	1	A43344	synaptic vesicle p
35	258	8.9	742	2	S27263	metabolite transpo
36	256	8.9	482	2	B69803	hypothetical metab
37	255.5	8.9	452	2	A64937	hypothetical prote
38	253	8.8	520	2	T23545	arabinose-proton s
39	250.5	8.7	472	2	S47089	hypothetical prote
40	250.5	8.7	517	2	T19962	sugar transporter
41	249.5	8.7	401	2	E69501	probable transport
42	248	8.6	416	2	H85786	metabolite transpo
43	243	8.4	461	2	D70073	hypothetical prote
44	242	8.4	312	2	T25852	metabolite transpo
45	241.5	8.4	457	2	E70070	metabolite transpo

ALIGNMENTS

RESULT 1
JW0089
organic cation transporter protein 2 - human
N:Alternate names: OCTN2
C:Species: Homo sapiens (man)
C:Date: 18-Jun-1998 #sequence_revision 10-Jul-1998 #text_change 28-May-1999
C:Accession: JW0089
R:Wu, X.; Prasad, P.D.; Leibach, F.H.; Ganapathy, V.
Biochem. Biophys. Res. Commun. 246, 589-595, 1998
A:Title: cDNA sequence, transport function, and genomic organization of human OCTN2,
A:Reference number: JW0089; MUID:98289574
A:Accession: JW0089
A:Molecule type: mRNA
A:Residues: 1-557 <WUA>
A:Cross-references: GB:AF057164; NID:g3273740; PIDN:AAC24828.1; PID:g3273741
A:Experimental source: placenta
C:Comment: This transporter functions in the elimination of cationic drugs and other

Query Match 100.0%; Score 2883; DB 2; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.8e-216;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRDYDEVTAFLGEMGFQRLIFLLSASII	PNGFTGLSSVFLIATPEHRCRVPDAA	LS 60
Db	1	MRDYDEVTAFLGEMGFQRLIFLLSASII	PNGFTGLSSVFLIATPEHRCRVPDAA	LS 60
QY	61	AWRNHTVPLRLRDGRVPHSCRRYRLAT	IANFSALGLEPGRDVLGQLESCLDGW	EFS 120
Db	61	AWRNHTVPLRLRDGRVPHSCRRYRLAT	IANFSALGLEPGRDVLGQLESCLDGW	EFS 120
QY	121	QDVYLSITVTENLVGECDDWKAPLTIS	LFVGVLGSTFSGQLSDFRGRKNVLF	VTMG 180
Db	121	QDVYLSITVTENLVGECDDWKAPLTIS	LFVGVLGSTFSGQLSDFRGRKNVLF	VTMG 180
QY	181	TGFSFLQIFSKNFEMVFLVVGMOISNY	AAVFLGTEILGKSVRIIFSLGVCIE	FA 240
Db	181	TGFSFLQIFSKNFEMVFLVVGMOISNY	AAVFLGTEILGKSVRIIFSLGVCIE	FA 240
QY	241	FGYVWLPLFAYFTRDWRMLLVALTMP	GVLCVALWFWIPESPRWLISQGRFE	AEVILRKA 300
Db	241	FGYVWLPLFAYFTRDWRMLLVALTMP	GVLCVALWFWIPESPRWLISQGRFE	AEVILRKA 300
QY	301	AKANGIVPSTIFDPSELQDLSSKKQ	QSHNILDLLRTWNIRMTIMSLMWT	ISVGYFG 360
Db	301	AKANGIVPSTIFDPSELQDLSSKKQ	QSHNILDLLRTWNIRMTIMSLMWT	ISVGYFG 360
QY	361	LSLDTPNLHGDIFVNCFLSAMVEVP	AYLAWLLQYLPRLRYSMATALFLG	SGVLLFMQLV 420
Db	361	LSLDTPNLHGDIFVNCFLSAMVEVP	AYLAWLLQYLPRLRYSMATALFLG	SGVLLFMQLV 420
QY	421	PPDLYYLATVLMVGKFGVTAAFSMV	YVYTAELYPVTVRNMGVGVSSSTA	SRGLSILSPYF 480

```

Db 421 PDLYLATLVVMGKFGVTAFAFVYVYTAELPTVVRNMGVGVSSATSLGSLSPYF 480
QY 481 VYLGAYDRFLPYLMGSLTILTAITLPLPSFGPLPDTIDQMLRVKGMKHKRTPSHR 540
Db 481 VYLGAYDRFLPYLMGSLTILTAITLPLPSFGPLPDTIDQMLRVKGMKHKRTPSHR 540
QY 541 MLKDGQERTILKSTAF 557
Db 541 MLKDGQERTILKSTAF 557

RESULT 2
JF0346
high-affinity carnitine transporter, Cnl - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: JF0346
R:Seikine, T.; Kusuohara, H.; Utsunomiya-Tate, N.; Tsuda, M.; Sugiyama, Y.; Kanai, Y.; Endo
Biochem. Biophys. Res. Commun. 251, 586-591, 1998
A:Title: Molecular cloning and characterization of high-affinity carnitine transporter
A:Reference number: JF0346; MUID:99011422
A:Accession: JF0346
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-557 <SEK>
A:Cross-references: DDBJ:AB017260; NID:g3869208; PID:BAA34399.1; PID:g3869209

Query Match 86.6%; Score 2496; DB 2; Length 557;
Best Local Similarity 85.3%; Pred. No. 7.8e-186;
Matches 475; Conservative 42; Mismatches 40; Indels 0; Gaps 0;

QY 1 MRDYDEVTAFLGEMGFORLIFFLLS-ASIIIPNGFTGLSSVFLIATPEHRCRVPDAANLSS 60
Db 1 MRDYDEVTAFLGEMGFORLIFFLLSASIIIPNGFTGLSSVFLIATPEHRCRVPDAANLSS 60
QY 61 AWRNHTVPLRLRDREVPHSCRRYRLATIANFSALGLEPGRDVLGQLEQESCLDGMWFS 120
Db 61 AWRNHSIPLETKDGRVQPSQRRYRLATIANFSALGLEPGRDVLGQLEQESCLDGMWFS 120
QY 121 QDVLSITVTENNVICEDDWKAPLTISLFFVGLVLLSGFISQSLSDRGKKNVFLVTGMQ 180
Db 121 QDVLSITVTENNVICEDDWKAPLTISLFFVGLVLLSGFISQSLSDRGKKNVFLVTGMQ 180
QY 181 TGFSLQIFSKNFEMFVFLVLMGQISNVAAAFVLGTEILGKSVRIIFSTLGCIFYA 240
Db 181 TGFSLQLFSVNFEMFVFLVLMGQISNVAAAFVLGTEILGKSVRIIFATLGCIFYA 240
QY 241 FGYVLPLEFAFIRDRMLLVALTMGPVLCVALLWFTPESPRWLISQGRPEAEVIIRKA 300
Db 241 FGYVLPLEFAFIRDRMLLVALTMGPVLCVALLWFTPESPRWLISQGRPEAEVIIRKA 300
QY 301 AKANGIVVPSTIFDPSELODLSKKQ--OSHNIIDLLRTNIRMTIMSLMWTISVGYG 360
Db 301 AKANGIVVPSTIFDPSELODLSKKQ--OSHNIIDLLRTNIRMTIMSLMWTISVGYG 360
QY 361 LSLDTPNLHGDIFFVNCFLSAMVEVPAYVLAALLQYLPRRYSMATLFLGGSVLLFQ 420
Db 361 LSLDTPNLHGDIFFVNCFLSAMVEVPAYVLAALLQYLPRRYSMATLFLGGSVLLFQ 420
QY 421 PDLYLATLVVMGKFGVTAFAFVYVYTAELPTVVRNMGVGVSSATSLGSLSPYF 480
Db 421 PSELYLATLVVMGKFGVTAFAFVYVYTAELPTVVRNMGVGVSSATSLGSLSPYF 480
QY 481 VYLGAYDRFLPYLMGSLTILTAITLPLPSFGPLPDTIDQMLRVKGMKHKRTPSHR 540
Db 481 VYLGAYDRFLPYLMGSLTILTAITLPLPSFGPLPDTIDQMLRVKGMKHKRTPSHR 540
QY 541 MLKDGQERTILKSTAF 557
Db 541 TOKDGGESPTILKSTAF 557

```

```

RESULT 3
T22509
hypoetical protein F52F12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22509
R:Matthews, L.
submitted to the EMBL Data Library, December 1996
A:Reference number: Z19573
A:Accession: T22509
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-576 <WIL>
A:Cross-references: EMBL:283228; PIDN:CA805732.1; GSPDB:GN00019; CESP:F52F12.1
A:Experimental source: clone F52F12
C:Genetics:
A:Gene: CESP:F52F12.1
A:Map position: 1
A:Introns: 16/1; 23/3; 50/1; 80/3; 108/2; 134/3; 221/1; 257/2; 319/3; 358/2; 490/3; 5

Query Match 27.0%; Score 777; DB 2; Length 576;
Best Local Similarity 31.4%; Pred. No. 1.5e-52;
Matches 165; Conservative 118; Mismatches 225; Indels 18; Gaps 6;

QY 3 DYDEVTAFLGEMGFORLIFFLLS-ASIIIPNGFTGLSSVFLIATPEHRCRVPDAANLSS 61
Db 16 DFDVLEQVNGYTYQIVFFFIICLTSLPSAFSAFNIFVVGNNPHTCHIPEGKEYLAP 75
QY 62 WRNHTVPLRLRDREVPHSCRRYRLATIANFSALGLEPGRDVLGQLEQESCLDGMWFSQ 121
Db 76 LTNDTQIL-----SCQYNETQINVFRAFTSAP-VDTYSDRISLVPQNGWDYDN 124
QY 122 DVLSITVTENNVICEDDWKAPLTISLFFVGLVLLSGFISQSLSDRGKKNVFLVTGMQ 181
Db 125 STYLSLVTETNLEFVLDQQAWEISTSTFYVGSFGICNLFYVADKFGRRSFVILTVLI 184
QY 182 GFSFLQIFSKNFEMFVFLVLMGQISNVAAAFVLGTEILGKSVRIIFSTLGCIFYAF 241
Db 185 VGTASSFAKDIESFIILRFTGLAFLQIPFIICMEFGMNSGR-IFSGLWTSLFFCA 243
QY 242 GYVLPLEFAFIRDRMLLVALTMGPVLCVALLWFTPESPRWLISQGRPEAEVIIRKAA 301
Db 244 AMALLGVAMFIRWRQLTFFCNAPFAFYIYFFFLPSRWSVSVGWADAKQKLLIA 303
QY 302 KANG---IVVPSTIFDPSELODLSKKQ--OSHNIIDLLRTNIRMTIMSLMWTISV 356
Db 304 KMGKSNVDDELVDMSKNHQAEEKTKRSHNVTDLFTPLNRKTLIVTYIWMNAI 363
QY 357 GYFGLSDTPNLHGDIFFVNCFLSAMVEVPAYVLAALLQYLPRRYSMATLFLGGSVLLF 416
Db 364 IYNGLTINSLNPVDDYWSFIINGAVELPGYFVWVLPLOCAGRRWTLAATMIVCGIGCVS 423
QY 417 MQLVPPDLYLATLVVMGKFGVTAFAFVYVYTAELPTVVRNMGVGVSSATSLGSL 476
Db 424 AMFMPDGYPLWVASAFIGKFGVSGFAVIYIFAGELYPTVVRNIGMGMSSVAGSGLL 483
QY 477 SPYFVYLGAYDRFLPYLMGSLTILTAITLPLPSFGPLPDTID 522
Db 484 APHIVNLGKIVKILPLLIMGLMALSAGILFTFLPETLGAFLPMTIE 529

RESULT 4
JC4884
organic cation transporter protein 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
C:Accession: JC4884
R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
Biochem. Biophys. Res. Commun. 224, 500-507, 1996
A:Title: cDNA cloning and functional expression of a novel rat kidney organic cation
A:Reference number: JC4884; MUID:96295517

```


A:Accession: JC4884

A:Molecule type: mRNA

A:Residues: 1-593 <ORU>

A:CROSS-references: DDBJ:DB3044; NID:g1502282; PIDN:BAAL1754.1; PID:d1012421; PID:g15022

A:Experimental source: kidney

C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 25.4%; Score 733; DB 2; Length 593;
Best Local Similarity 35.2%; Pred. No. 3.9e-49;
Matches 198; Conservative 89; Mismatches 234; Indels 42; Gaps 15;

QY 1 MRDYDEVTAFLEGWGPQRLIFELLSAIIIPNGFTGL--SSVPLIATPEHRCRVPDAANL 58

Db 1 MSTVDDILEHIGFHLFKQKTFELL--ALLSGAFTPIYGVIFELGTPDHCWSPGAANKL 58

QY 59 SS--AWR-----NHTVP-LRLRGRVPHSCRRYR-----LATIANFSALGLEPGRDV 103

Db 59 SQRGWSQAEELNVTVPGLGPSDEASFLSQCMRYEVDWQNSTLDCVDPSSLA-----A 112

QY 104 DLGOLESCLDGWESQDYVLTIVTEWNLVCEDDWKAPLITSLFVGVLLGSFISGQL 163

Db 113 DRNPLPGCPGHEGVVNTPG--SSIVTEFNLCVCAHSMWMLDLFQSVVNVGFFIGAMMIGYL 170

QY 164 SDRGRKNLVFTVMGQTGFSLQIFSKNFEMFVFLVFLVGMGQISNYVAAFVLGTEILG 223

Db 171 ADRGRKFCLLVTLINAIISGALMAISPNYAWMLVRFELGLVSKAGWLGILITLITFVG 230

QY 224 KSVRIIFSTLGVCIYAF--GYMVLPLFAFYFIRDWRMLLVALTMPGVLCVALWWEIPESP 281

Db 231 LGYR---RMVGICYQTAFTVGLILLAGVAVIENWRLOQFAVTLNFCFLYFWCIPESP 287

QY 282 RWLSQGRFEEAVIIRKAKANGIYVVPSTIFDPSELQDLSSKKQOSHNLIDLLRTWNIR 341

Db 288 RWLSQNKIVKAMKIHKIAKNGKSPVSLQNLTPDEDAGKLPKPS--ILDLVPTQIR 345

QY 342 MVTIMSIMLAWTISVGVFGLSLDTPNLHGD-IFVNCFLSAMVEVPAYVLAWLLQLVLP 400

Db 346 KHTLILMYNFTSSVLYQGLIMHM-GLAGDNIYLDFFYSALVEFPAAFTIITIDRVGR 404

QY 401 YSMATALFLGGSVLLFMQLVPPDLYLATVLMVVGKFGVTAAPSMVYVTAELYPTVVRN 460

Db 405 YPNAVSNMVAGAACLASVFIPDDLOMKITIACLGRMGITMAYEMVCIYNAELYPTIYRN 464

QY 461 MGVGVSSTARLGSILSPFYV-LGAYDRFLPYILMGSLTITAILTLPLPESFGTLPD 519

Db 465 LGVLVCSMCDIGGIITPFLVRLTDIWMEEFLVFAVAVGLVAGALVLLLPETKGAALPE 524

QY 520 TID-----QMLRVKGMKHKRTPS 537

Db 525 TIEDAENMQRPKRKKNLPPS 547

RESULT 5

S50862

organic cation transport protein Ocr1 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Nov-1999

C:Accession: S50862; S78533; I58089

R:Gruendemann, D.; Gorboulev, V.; Gambaryan, S.; Veyhl, M.; Koepsell, H.

Nature 372, 549-552, 1994

A:title: Drug excretion mediated by a new prototype of polyspecific transporter.

A:Reference number: I58089; MUID:95082907

A:Accession: S50862

A:Molecule type: mRNA

A:Residues: 1-556 <GRU>

A:CROSS-references: EMBL:X78855; NID:g633621; PIDN:CAA55411.1; PID:g633622

A:Experimental source: kidney

R:Gorboulev, V.G.

submitted to the EMBL Data Library, January 1995

A:Reference number: S78533

A:Accession: S78533

A:Molecule type: mRNA

A:Residues: 1-342, 'N', 344-556 <GOR>

A:CROSS-references: EMBL:X78855; NID:g633621; PIDN:CAA55411.1; PID:g633622

C:Keywords: glycoprotein; phosphoprotein; transmembrane protein

F:20-46/Domain: transmembrane #status predicted <TM1>

F:154-171/Domain: transmembrane #status predicted <TM2>

F:178-197/Domain: transmembrane #status predicted <TM3>

F:243-260/Domain: transmembrane #status predicted <TM4>

F:267-283/Domain: transmembrane #status predicted <TM5>

F:350-366/Domain: transmembrane #status predicted <TM6>

F:380-398/Domain: transmembrane #status predicted <TM7>

F:406-425/Domain: transmembrane #status predicted <TM8>

F:435-452/Domain: transmembrane #status predicted <TM9>

F:469-485/Domain: transmembrane #status predicted <TM10>

F:494-514/Domain: transmembrane #status predicted <TM11>

F:71.97.113.432/Binding site: carboxylate (Asn) (covalent) #status predicted

F:286.292/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pred

F:296.343.550/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status

Query Match 24.7%; Score 712; DB 2; Length 556;

Best Local Similarity 33.8%; Pred. No. 1.5e-47;

Matches 193; Conservative 96; Mismatches 230; Indels 52; Gaps 14;

QY 1 MRDYDEVTAFLEGWGPQRLIFELL---SASIIIPNGFTGLSSVFLIATPEHRCRVPDAAN 57

Db 1 MPTVDVLEQVGEFGWFOQAFLLLCLISASLAPI-YVGI--VFLGFTPGHYCQNGVAE 57

QY 58 LSS--AWR-----NHTVP-LRLRGRVPHSCRRYRL-----ATTANFSAL 95

Db 58 USQRGWSQAEELNVTVPGLGPSDEASFLSQCMRYEVDWQNSTLDCVDPSSLVANKSQL 117

QY 96 GLEPGRDVLGOLESCLDGWESQDYVLTIVTEWNLVCEDDWKAPLITSLFVGVLL 155

Db 118 PLGP-----CEHGMY--DTPGSSIVTEFNLCVGDANKVDLDFOSCVNLGFFL 162

QY 156 GSFTSGQLSDRGRKNLVFTVMGQTGFSLQIFSKNFEMFVFLVFLVGMGQISNYVAAF 215

Db 163 GSLVGVYIADRGRKLCLLVTLVTSVSGVLTAVAPDYSMLLFRLLQGMVSKGWSVSGY 222

QY 216 VLGTIELKSKVRIIFSTLGVCIYAF--GYMVLPLFAFYFIRDWRMLLVALTMPGVLCVAL 273

Db 223 TLITEFVSGYR---RTTAYIQMAFTVGLVGLAGVATAIPDRWMLQLAVSLFTFLLY 279

QY 274 WMFIPESPRWLISQGRFEEAVIIRKAKANGIYVVPSTIFDPSELQDLSSKKQOSHNLID 333

Db 280 YWFVPEPRWLLSQKRTTRAVIMEQIAQKNGKVPDPADLKMCLLEEDASER--SPSEAD 337

QY 334 LLRTWNIRMTIMSLMWTISVGVFGLSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLL 393

Db 338 LFRTPTLRKHTVILMYLWFSCAVLYQGLIMHVGTAGANLYLDFFYSILVEPPAAFIILVT 397

QY 394 LOYLPRRYSMATALFLGGSVLLFMQLVPPDLYLATVLMVVGKFGVTAAPSMVYVTAEL 453

Db 398 IDRTGRITYPIAASNLVTGAACLLMIFIPHEUHLNVTIACLRMGATIVLQMVCLVNAEL 457

QY 454 YPTVVRNMGVGSSTASRLGSILSPFYV-LGAYDRFLPYILMGSLTITAILTLPLFLPES 512

Db 458 YPTFIRNLGMVCSALCDLGGIFTPFEMVRLMEVWQALPLILFGLVGLITAGAMTLLLPET 517

QY 513 FGTPLPDTIDQMLRVKGMKHKRTPSFRMLK 543

Db 518 KGVALPETIEEAENL-GRKSKAKENTIIYLQ 547

RESULT 6

T27870

hypotheical protein ZK455.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T27870

R:White, S.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z20432

A;Map position: 1

Query Match	18.1%; Score 522; DB 2; Length 527;
Best Local Similarity	28.3%; Pred. No. 7.7e-33;
Matches 155; Conservative 100; Mismatches 206; Indels 86; Gaps 17;	
Qy 5	DEVTAFTLGEMGPFORLFFLLSASIIIPNGTGTGSSVFLIATPEHRCRVPDAAANLSSAWRN 64
Db 22	DTIESYIGSGFQAQALVAFSGVFDQAQTIFSVTDEPTDTHWC----- 67
Qy 65	HTVPLRLRDGREVPHSCRRYRLATIANFSALEGPGRVDLGOLEQESCLDGEFSDVY 124
Db 68	-----TDSNSICHE-----SISNICIL-----PKTAWSDYSPHV- 97
Qy 125	LSTIVTNLWVCEDDKAPLTISLFFVGVLLGSPISQLSD-RFRGNKVLFTVMGOTGF 183
Db 98	--SWISEMGLQACGSFKVGLPESFFVCGIGIGLVLTADSLGRKNMLFLSCLVNAIS 155
Qy 184	SFLQITKSNFEMFVFLVGLVMGQISNKVAAFVLGTILGSKVRIIPSTLGVCIYAF-- 241
Db 156	TMLVFSNINWYAVLRFVNGFGRATICTCALVSTELVCKKWR--GRVGIMSFGEFML 212
Qy 242	GYMVLPLPAYIR--DWRMLLVALTMGVL-CVALWFIPIESPRLWISQCRFEAEVIR 298
Db 213	GLSLPLMAYMNRGSSWRILYAWTSIPTIYCVLVRFCVCEPRKLVFVRGRREASILK 272
Qy 299	KAARANGTVVPST-----TFDSEIQDLSKKQOQSHNILDLLRT-----WNT-RM 342
Db 273	RVAS-----IFSTDVSSGGAISMFSFSLPPEDEEKPTNVNIFTMKVLVEKRWALKRL 327
Qy 343	VTIMSIMLWMTIVSGYFGLSLDTNPLHGDIIVNCFLSAWEVPAYVLAWLLQYLPREYS 402
Db 328	SAVMAIA--FGIGLVYVGMPLALNLFNYLSAANLMDLPANLITFLVDKLSRNA 385
Qy 403	NATALFLGG--SVLLF----MQLVPPDLYLATVLMVKGFGVTAAFSMVYVYTAELPYT 456
Db 386	LIGETALGVSSVLIFALHNHRIGNHGALQLA--LELISYSFACSFAFMEMIYITIELFPT 443
Qy 457	VYRNMGVSVSTASRLGSLSPYFYLIGAYDRFLPYILMGSLTILTAILFLFPESCTP 516
Db 444	CYRNSAIAARQALVIGGVFSPIVMYAGRKNAFWSFGLGLAIGLLGLFAVLGFPETRGSD 503
Qy 517	LPDITDQ 523
Db 504	LCDTWDE 510

RESULT 9
 C96758
 probable protein transporter T18K17.11 [imported] - Arabidopsis thaliana
 C96758
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96758
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ausen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: C96758
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-539 <STO>
 A:Cross-references: GB:AE005173; NID:g6598860; PIDN:AAF18714.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T18K17.11
 A:Map position: 1

[illegible]

RESULT 10
B86299
hypothetical protein AAD34691.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B86299
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 815-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: B86299
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-518 <STC>
A:Cross-references: GB:AF005172; NID:g4966360; PIDN:AAD34691.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match	16.3%	Score 470.5	DB 2	Length 518
Best Local Similarity	28.2%	Pred. No. 7.4e-29		
Matches	136	Conservative 94	Mismatches 218	Indels 35
Gaps	12			
QY	55	AANLSSAW---RNHTVPLRLRDGVEPHSCRRYRLATIANFSAIGLEPGRDVDLQGLEOE	111	
DB	44	AALVSFAWFFDAQOQTFVTDSOPMHCNDSRDVSCVNTSSNL-----	89	
QY	112	SCL---DQWFSQDYLSTIVTENNLVCEDDKAPLTISLFFVGLVSGFISGQSLD-RF	167	
DB	90	-CTLPNQWSDNLNPHYS-IISENGLOACAGFLKGPFPASSFFLGLGIALSTLADSSL	147	
QY	168	GRKNVLPVTHMGQTFSEFLQIFSKNEFMFVFLVLMGMOISNYAAAFVLGTEILGKSVR	227	
DB	148	GRKNMLLSLIMSLSMLTAFSTISWYAFRLFLNGCGRATIGTCALVLSTELVGKWR	207	
QY	228	IIFSTLGCIFYAP--GYMWLPPLFAFYI--RDRMMLLVALTMPGVL-CVALWAFIPESPR	282	
DB	208	--GOVGAMGFFCTLGLSLPLMLGYINEGNSRNLYVVTSPITLIYCCLVRSFVRESPR	264	
QY	283	WLISQRFEEAEVIRKAAKANGIVVPSTIFDSELODLSKKQKQSHNILDLL--RTWNI	340	
DB	265	WLIYVGRKEEAVSILQIA-SNAITMSFTNLCPEVDNQSKSPNDYDALKILVRKWSF	323	
QY	341	RMVIMSTMLMWTISVGYFGLSLDTNMLHGDIFVNCFLSAMVEVPAYLAWLLQLYLR	400	
DB	324	RRL-LAAMVVGFGIMYVYGMPLALNPLNPLNLGVFNALSEFFAFLITFFIDKINRR	382	
QY	401	YSMATALFLGSGVLLFMQVPPDLYLAVLMVGRFGVTAAFSMYVYVTAELYPTVVRN	460	
DB	383	DALIGFTALSGISALIAVLGQOLGSLQIVLELVSPFSACTAFNMTLIYTIEMFPTCVRN	442	
QY	461	MGVSVSTASRLGSLSPYVYLGAVDRLPYLTMGSLTILTAILTFLFPESFGPLPDT	520	
DB	443	SAISMVQALVFGGVSPVMAAGRENQWSGLFGLIIGLGLFVGLPGRGSLVCDT	502	
QY	521	IDQ 523		
DB	503	MDE 505		
RESULT	11			
T21682				
hypothetical protein F32H5.4 - Caenorhabditis elegans				
C:Species: Caenorhabditis elegans				
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999				
C:Accession: T21682				
R:Steward, C.				
submitted to the EMBL Data Library, November 1996				
A:Reference number: Z19458				
A:Accession: T21682				
A:Status: preliminary; translated from GB/EMBL/DBDJ				
A:Molecule type: DNA				
A:Residues: 1-528 <WIL>				
A:Cross-references: EMBL:Z81524; PIDN:CAB04250.1; GSPDB:GNO0023; CESP:F32H5.4				
A:Experimental source: clone F32H5				
C:Genetics:				
A:Gene: CESP:F32H5.4				
A:Map position: 5				
A:Introns: 63/1; 108/3; 148/3; 249/2; 408/1; 472/3				
Query Match	14.3%	Score 411	DB 2	Length 528
Best Local Similarity	26.4%	Pred. No. 3.1e-24		
Matches	126	Conservative 91	Mismatches 189	Indels 92
Gaps	18			
QY	100	GRVDLQGLEQSCLDGWEF-----SDVYLSTIVTENNLVCEDDKAPLTISL	149	
DB	73	GPDVFI---KDICKSFEFHNPPNCTITSKYDEF--SINVDYGHFCGEGAWVKTISVQ	126	
QY	150	FVGLVSGFISGDSRFRGKNVLFVTHMGQTFSEFLQIFSKNEFMFVFLVLMG--G	206	
DB	127	MVGLIGSVTSAGVADRYGLKVLSCFVWSSILNTFAKDLIYTIINTLISIFKGG	186	

submitted to the EMBL Data Library, April 1993
 A:Reference number: Z19704
 A:Accession: T23190
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-529 <WIL>
 A:Cross-references: EMBL:Z22175; PIDN:CAA80131.1; GSPDB:GN00021; CESP:ZK637.1
 A:Experimental source: clone K01f9
 C:Genetics:
 A:Gene: CESP:ZK637.1
 A:Map position: 3
 A:Introns: 31/2; 121/3; 150/3; 192/2; 255/3; 422/2; 472/3
 C:Superfamily: Caenorhabditis elegans glucose transport protein

Query Match	12.3%	Score 355.5;	DB 2;	Length 529;
Best Local Similarity	26.5%	Pred. No. 6.3e-20;		
Matches 140;	Conservative 97;	Mismatches 186;	Indels 105;	Gaps 24;

QY	52	VPOAANLSSAARNHIVPLRLRDGREVPHSCRRYRLATIANFSALGLEP--GRDVD-----	104
DB	10	VLEASNLEAYVDLTAKQLIKEIRHVGDDE-AVRYSNLDDRTLEG-EPTDQRSPPDSEKTF	67
QY	105	-----LGQLEQE-SCLDGWFSQD-----VYLSTI-----VTEWNLVCEDDWKAPLT	145
DB	68	TVDEAVEALGFGRFQKLKSLITGMAWMAAMENMLLSLSPALACEWGI---SSVQQAIV	124
QY	146	ISLFFVGVLLGSLGSDREG-RKNVLEVT-----MGMTQGS---FLQIFSKNFEMP	196
DB	125	TTCVFSGMMLSTFNKICDQFGRRKGLTFSTLVACIMGVISGMSPHFYLLFFRGLTGF	184
QY	197	VVLFVLVGMGOISNYAAV--LGTEILGKSVRIIFSTLGVCFYAFGYMVLPLFAYFIR	254
DB	185	-----GIGGVPSVTLVAEFLPTAQRACVVLIES-----FWAIGAVFEALLAYFVM	231
QY	255	D---WRMLLVALTMP-CVLCVALWFIPESPRWLISQGRFEAEVIRKAANGIVVPS	310
DB	232	ESFGWALMFLSLPLIGIFAVASFW-LPESARFDMASGHPERALETQAAARNRVQLPT	290
QY	311	TIPDPSELQDLSSKKQOOSHNLDL--LRTWNIWMVTIMSLWMTISVGYGLSLDTPNL	368
DB	291	G-----RLVSTKAGSESRGDIANLLSPDLRKTILLWCIAWATAFSYGWLFTTVL	343
QY	369	-----HGDIFVN-----CFLSAMVEVPAYVLAWLLQLYLPRRYSMA	404
DB	344	FQSHDECHGGLFSNGTQMEVCOPLTRSDYFDLLSTLAEFPGLIITVLIIEWFGKTKMA	403
QY	405	T--ALFLGGSVLLFQVLPDLYLATVLMYKFGVTAAFSMWVYVYTAELYPTVVRNMG	462
DB	404	LEVAVFAITFLLYICLD-----RFTVTVLIFVARAFISGAFOCAYVYTPVYPTTLRAVG	459
QY	463	VGVSTASRLGSLRPYFVYLGAYDRFLPYILMGSLTITLITLFLP	510
DB	460	LGTCSAMARIGAIVA-----SEKSLSLPIGIYGTAAIILGLIASLSLP	501

Search completed: January 25, 2002, 10:09:48
 Job time: 84 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:15:19 ; Search time 18.96 Seconds

(without alignments)
1077.126 Million cell updates/sec

Title: US-09-521-195-3

Perfect score: 2883

Sequence: 1 MRDYDEVTAFLGEGWPFQRL.....HTRMLKDGQRPILKSTAF 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2883	100.0	557	1	OCN2_HUMAN
2	2513	87.2	557	1	OCN2_MOUSE
3	2509	87.0	557	1	OCN2_RAT
4	387.5	13.4	751	1	YLX5_CAEEL
5	357	12.4	1222	1	YMP3_CAEEL
6	355.5	12.3	529	1	YUUL_CAEEL
7	308	10.7	445	1	YGCS_ECOLI
8	291.5	10.1	435	1	YT13_CAEEL
9	287	10.0	400	1	YCEI_BACSU
10	267	9.3	443	1	YAAU_ECOLI
11	265	9.2	459	1	YDUK_ECOLI
12	262.5	9.1	742	1	SYV2_RAT
13	256	8.9	482	1	YFIG_BACSU
14	255.5	8.9	452	1	YDJE_ECOLI
15	252.5	8.8	457	1	PCAK_ACICA
16	250.5	8.7	472	1	ARAE_KLEOX
17	243.5	8.4	551	1	HGT1_KLULA
18	241.5	8.4	472	1	ARAE_ECOLI
19	240.5	8.3	495	1	GTR3_CANFA
20	239.5	8.3	451	1	YFAJ_BACSU
21	239	8.3	461	1	CSBC_BACSU
22	237.5	8.2	496	1	GTR3_HUMAN
23	234	8.1	494	1	GTR3_SHEEP
24	227	7.9	490	1	GTRL_CHICK
25	226.5	7.9	484	1	GALP_ECOLI
26	224.5	7.8	413	1	MUCK_ACICA
27	223	7.7	451	1	GTRL_PIG
28	222	7.7	763	1	RG2T_YEAST
29	221	7.7	466	1	BENK_ACICA
30	221	7.7	492	1	GTRL_RAT
31	220	7.6	492	1	GTRL_BOVIN
32	219.5	7.6	491	1	XYLE_ECOLI
33	219	7.6	493	1	GTR3_MOUSE

RESULT 1

ID	OCN2_HUMAN	STANDARD;	PRT;	557 AA.
AC	076082;			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	ORGANIC CATION/CARNITINE TRANSPORTER 2 (SOLUTE CARRIER FAMILY 22, MEMBER 5) (HIGH-AFFINITY SODIUM-DEPENDENT CARNITINE COTRANSPORTER).			
GN	SLC22A5 OR OCTN2			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=9828574; PubMed=9618255;			
RA	Wu X., Prasad P.D., Leibach F.H., Ganapathy V.;			
RT	"cDNA sequence, transport function, and genomic organization of human OCTN2, a new member of the organic cation transporter family.";			
RL	J. Biochem. Biophys. Res. Commun. 246:589-595(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=98352077; PubMed=9685390;			
RA	Tamai I., Ohashi R., Nezu J.-I., Yabuuchi H., Oku A., Shimane M., Sai Y., Tsuji A.;			
RT	"Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter OCTN2.";			
RL	J. Biol. Chem. 273:20378-20382(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99113835; PubMed=9916797;			
RA	Nezu J., Tamai I., Oku A., Ohashi R., Yabuuchi H., Hashimoto N., Nikaide H., Sai Y., Koizumi A., Shoji Y., Takada G., Matsuishi T., Yashino M., Kato H., Ohura T., Tsujimoto G., Hayakawa J., Shimane M., Tsuji A.;			
RT	"Primary systemic carnitine deficiency is caused by mutations in a gene encoding sodium ion-dependent carnitine transporter.";			
RL	Nat. Genet. 21:91-94(1999).			
RN	[4]			
RP	CHARACTERIZATION.			
RX	MEDLINE=99384224; PubMed=10454528;			
RA	Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H., Chen J., Conway S.J., Ganapathy V.;			
RT	"Functional characteristics and tissue distribution pattern of organic cation transporter 2 (OCTN2), an organic cation/carnitine transporter.";			
RL	J. Pharmacol. Exp. Ther. 290:1482-1492(1999).			
RN	[5]			
RP	VARIANT CDSP GLN-169.			
RX	MEDLINE=99355597; PubMed=10425211;			
RA	Burwinkel B., Kreuder J., Schweitzer S., Vorgerd M., Gempel K., Gerbitz K.-D., Kilian M.W.;			
RT	"Carnitine transporter OCTN2 mutations in systemic primary carnitine deficiency: a novel Arg169Gln mutation and a recurrent Arg282ter			

P17809 mus musculus
P28568 gallus galli
P11166 homo sapien
P13355 oryctolagus
P53142 saccharomyc
Q51955 pseudomonas
P10870 saccharomyc
Q07647 rattus norv
P40895 saccharomyc
P54862 saccharomyc
Q01440 leishmania
P38695 saccharomyc

ALIGNMENTS

|||||
Db 481 VYLGADYRFLPYLMGSLILTRAILFLPESFGTLPDTIDQMLRVKGNKRPSTHR 540
QY 541 MLKDQGERPTILKSTAF 557
|||||
Db 541 MLKDQGERPTILKSTAF 557
RESULT 2
OCN2_MOUSE STANDARD; PRT; 557 AA.
AC Q92088;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE ORGANIC CATION/CARNITINE TRANSPORTER 2 (SOLUTE CARRIER FAMILY 22,
DE MEMBER 5) (HIGH-AFFINITY SODIUM-DEPENDENT CARNITINE COTRANSPORTER).
GN SLC22A5 OR OCTN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=99113835; PubMed=9916797;
RA Nezu J., Tamai I., Oku A., Ohashi R., Yabuuchi H., Hashimoto N.,
RA Nikaido H., Sai Y., Koizumi A., Shoji Y., Takada G., Matsuishi T.,
RA Yashino M., Kato H., Ohura T., Tsujimoto G., Hayakawa J., Shimane M.,
RA Tsuji A.;
RT "Primary systemic carnitine deficiency is caused by mutations in a
RT gene encoding sodium ion-dependent carnitine transporter.";
RT Nat. Genet. 21:91-94(1999).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT JVS ARG-352.
RC STRAIN=C3H;
RX MEDLINE=99057546; PubMed=9837751;
RA Lu K., Nishimori H., Nakamura Y., Shima K., Kuwajima M.;
RT "A missense mutation of mouse OCTN2, a sodium-dependent carnitine
RT cotransporter, in the juvenile visceral steatosis mouse.";
RL Biochem. Biophys. Res. Commun. 252:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=99384224; PubMed=10454528;
RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
RA Chen J., Conway S.J., Ganapathy V.;
RT "Functional characteristics and tissue distribution pattern of organic
RT cation transporter 2 (OCTN2), an organic cation/carnitine
RT transporter";
RL J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
CC -|- FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE
CC TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
CC INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
CC CARNITINE.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -|- DISEASE: DEFECTS IN SLC22A5 ARE THE CAUSE OF JUVENILE VISCERAL
CC STEATOSIS (JVS).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
CC CATION SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB015800; BAA36590.1; -;
DR EMBL; AF111425; AAC99787.1; -;
DR EMBL; AF110417; AAD54060.1; -;
DR MGI; MGI:1329012; SLC22a5.
DR InterPro; IPR003662; sub_transporter.

Pfam: PF00083; sugar_tr: 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1: 1. Disease mutation.
KW Transport; Transmembrane; Glycoprotein; POTENTIAL.
FT TRANSMEM 21 41
FT TRANSMEM 143 163
FT TRANSMEM 173 193
FT TRANSMEM 198 218
FT TRANSMEM 233 253
FT TRANSMEM 258 278
FT TRANSMEM 342 362
FT TRANSMEM 374 394
FT TRANSMEM 407 427
FT TRANSMEM 431 451
FT TRANSMEM 489 509
FT CARBOHYD 57 57
FT CARBOHYD 64 64
FT CARBOHYD 91 91
FT CARBOHYD 322 322
FT VARIANT 352 352
SQ SEQUENCE 557 AA; 62779 MW; 6093F0EE9612B204 CRC64;
Query Match 87.2%; Score 2513; DB 1; Length 557;
Best Local Similarity 85.5%; Pred. No. 1.1e-154;
Matches 476; Conservative 39; Mismatches 42; Indels 0; Gaps 0;
QY 1 MRDYDEVTAFTGEGWPPFORLIFFLLSASIIPINGFTGLSSVFLIATPEHRCRVPDAANLSS 60
|||||
Db 1 MRDYDEVTAFTGEGWPPFORLIFFLLSASIIPINGFTGLSSVFLIATPEHRCRVPDAANLSS 60
QY 61 AWRNHTVPLRLRDGREGVPHSCRRYRLATIANFSALEGPGRDVLQLEQESCLDGEWFEFS 120
|||||
Db 61 AWRNHSIPLTKDGRQVQPKCRRYRLATIANFSELGLEPGRDVLQLEQESCLDGEWFEFS 120
QY 121 QDYVLTSTIVTENLYCEDDWKAPLITISLFFVGLVLSFISQGLSDRFRGNKLVFTMGMQ 180
|||||
Db 121 KDVELSTIVTEWDLVCKDDWKAPLITISLFFVGLVLSFISQGLSDRFRGNKLVFTMGMQ 180
QY 181 TGFSELOFSKNFEMFVFLVGLVGMGOISNVAAVFLTEILGKSVRIIFSTLGVCFIYA 240
|||||
Db 181 TGFSELOFSKNFEMFVFLVGLVGMGOISNVAAVFLTEILGKSVRIIFSTLGVCFIYA 240
QY 241 FGMYVLPLFAYEIRDRWMLLVALTMPGVLVCLVAFIPESPRWLISQGRFEEAEVIRKA 300
|||||
Db 241 FGFVLPFAYEIRDRWMLLVALTMPGVLVCLVAFIPESPRWLISQGRFEEAEVIRKA 300
QY 301 AKANGIVVPSTIFDPSLODLSKKQKSHNLDLRTWNIRMTIMSLMWTISVGVYFG 360
|||||
Db 301 AKINGIVAPSTIFDPSLODLSKKQKSHNLDLRTWNIRMTIMSLMWTISVGVYFG 360
QY 361 LSLDTPNLHGDIIVNCFELSAMVEPAYVLAALLQYLPERRYSMATALELGGSVLLFMOLV 420
|||||
Db 361 LSLDTPNLHGDIIVNCFELSAMVEPAYVLAALLQYLPERRYSMATALELGGSVLLFMOLV 420
QY 421 PDLXPLATVLMVGKFGVTAFAFSMVYVYTAELYPTVVRNMGVSVSTASRLGSLSPYF 480
|||||
Db 421 PSELYFLSTALVMVGKFGITSAYSVMYVYTAELYPTVVRNMGVSVSTASRLGSLSPYF 480
QY 481 VYLGADYRFLPYLMGSLITLITLTLFPESEFGTLPDTIDQMLRVKGNKRPSTHR 540
|||||
Db 481 VYLGADYRFLPYLMGSLITLITLTLFPESEFGTLPDTIDQMLRVKGNKRPSTHR 540
QY 541 MLKDQGERPTILKSTAF 557
Db 541 MORDGEESTVLKSTAF 557
RESULT 3
OCN2_RAT STANDARD; PRT; 557 AA.
ID OCN2_RAT
AC Q70594; Q90WLO;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)

20-AUG-2001 (Rel. 40, Last annotation update)
 ORGANIC CATION/CARNITINE TRANSPORTER 2 (SOLUTE CARRIER FAMILY 22,
 MEMBER 5) (HIGH-AFFINITY SODIUM-DEPENDENT CARNITINE COTRANSPORTER)
 (UST2R) (CT1).
 GN SLC22A5 OR OCTN2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney.
 RX MEDLINE=98200080; PubMed=9541011;
 RA Schoemig E., Spitzberger F., Engelhardt M., Martel F., Oerding N.,
 RA Gruendemann D.;
 RT "Molecular cloning and characterization of two novel transport
 RT proteins from rat kidney.";
 RL FEBS Lett. 425:79-86(1998).
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Intestine;
 RX MEDLINE=99011422; PubMed=9792817;
 RA Sekine T., Kusahara H., Utsunomiya-Tate N., Tsuda M., Sugiyama Y.,
 RA Kanai Y., Endou H.;
 RT "Molecular cloning and characterization of high-affinity carnitine
 RT transporter from rat intestine.";
 RL Biochem. Biophys. Res. Commun. 251:586-591(1998).
 RN
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX MEDLINE=99384224; PubMed=10454528;
 RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
 RA Chen J., Conway S.J., Ganapathy V.;
 RT "Functional characteristics and tissue distribution pattern of organic
 RT cation transporter 2 (OCTN2), an organic cation/carnitine
 RT transporter.";
 RL J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
 CC
 CC -1- FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE
 CC TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
 CC INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
 CC CARNITINE.
 CC
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE PROXIMAL AND DISTAL TUBULES
 CC AND IN THE GLOMERULI IN THE KIDNEY, IN THE MYOCARDIUM, VALVES, AND
 CC ARTERIOLES IN THE HEART, IN THE LABYRINTHINE LAYER OF THE
 CC PLACENTA, AND IN THE CORTEX, HIPPOCAMPUS, AND CEREBELLUM IN THE
 CC BRAIN.
 CC
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
 CC CATION SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AJ001933; CAA05106.1;
 CC EMBL; AB017260; BAA34399.1;
 CC EMBL; AF110416; AAD54059.1;
 CC InterPro: IPR003662 sub_transport.
 CC Pfam: PF00083; sugar_tr; 1.
 CC PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 KW Transport; Transmembrane; Glycoprotein.
 FT TRANSMEM 21 41
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 198 218 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 258 278 POTENTIAL.
 FT TRANSMEM 342 362 POTENTIAL.
 FT TRANSMEM 374 394 POTENTIAL.
 FT TRANSMEM 407 427 POTENTIAL.

FT TRANSMEM 431 451 POTENTIAL.
 FT TRANSMEM 489 509 POTENTIAL.
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 275 275 W -> G (IN REF. 2).
 SQ SEQUENCE 557 AA; 62567 MW; 172472E7B0A5F043 CRC64;
 Query Match 87.0%; Score 2509; DB 1; Length 557;
 Best Local Similarity 85.5%; Pred. No. 2e-154;
 Matches 476; Conservative 42; Mismatches 39; Indels 0; Gaps 0;
 QY 1 MRDYDEVTAFLGEGWPFQRLIFLLSASIPNGFTGLSSVFLIATPEHRCRVDPDAANLSS 60
 DB 1 MRDYDEVTAFLGEGWPFQRLIFLLSASIPNGFTGLSSVFLIATPEHRCRVDPDAANLSS 60
 QY 61 AWRNHTVPLRLDRGREVPHSCRRYRLATIANFALGEPGRDVLQLEQECIDGWEFS 120
 DB 61 AWRNHSIPLETKDGRQVPCRRYRLATIANFALGEPGRDVLQLEQECIDGWEYN 120
 QY 121 ODVYLSITVTEWNLVCEDDWKAPLITSLFPVGLLSGFTSGQLSDRFGKKNVLFVTGMQ 180
 DB 121 KDVELSTIVTEWDLVKDDWKAPLITSLFVGVLMGFTSGQLSDRFGKKNVLFVTGMQ 180
 QY 181 TGSFLOIFSKNFEMFVFLVGMGOISNYVAAFVLGTEILGKSVRIIFSTLGVCIYA 240
 DB 181 TGSFLOIFSVNFEMFVFLVGMGOISNYVAAFVLGTEILGKSVRIIFSTLGVCIYA 240
 QY 241 FGVNVLPLFAFIRDMRLMLVALTMPGVLCVLAWMFTIPESPRWLISQGRFEEAEVIRKA 300
 DB 241 FGVNVLPLFAFIRDMRLMLVALTMPGVLCVLAWMFTIPESPRWLISQGRFEEAEVIRKA 300
 QY 301 AKANGIVPSTIFDPSLQDLSSKKQSHNILLDTNIRWVIMSLMWTISVGVYFG 360
 DB 301 AKANGIVPSTIFDPSLQDLSSKKQSHNILLDTNIRWVIMSLMWTISVGVYFG 360
 QY 361 LSDDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOVLPYRYSMATLFLGGSVLLFPMQLV 420
 DB 361 LSDDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOVLPYRYSMATLFLGGSVLLFPMQLV 420
 QY 421 PDLIYLAIVLVVGVKFGVTAAFSMVYVYTAELYPVVRNMVGVSSTASRLGSLSPYF 480
 DB 421 PSELYLSTALVVMVGFGITSAISVMVYVYTAELYPVVRNMVGVSSTASRLGSLSPYF 480
 QY 481 VYLGAYDRELPLILMGSLTILTAITLFLPESFGTLPDITDQMLRVKGMKHKTPSHFR 540
 DB 481 VYLGAYDRELPLILMGSLTILTAITLFLPESFGTLPDITDQMLRVKGMKHKTPSHFR 540
 QY 541 MLKDGQERPTILKSTAF 557
 DB 541 TQKDGESPTVLKSTAF 557
 RESULT 4
 YLX5_CAEBL STANDARD; PRT; 751 AA.
 AC P46501;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHETICAL 84.8 KDA PROTEIN F23F12.5 IN CHROMOSOME III.
 GN F23F12.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2.
 RA Du Z.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U12965; AAA20607.1; -.
DR WormPep; F23F12.5; CE01252.
DR InterPro; IPR002184; Srb.
DR InterPro; IPR003662; sub_trnsporttr.
DR Pfam; PF02175; Srb; 1.
DR Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 57 77 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 137 157 POTENTIAL.
FT TRANSMEM 200 220 POTENTIAL.
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
FT TRANSMEM 379 399 POTENTIAL.
FT TRANSMEM 410 430 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 515 535 POTENTIAL.
FT TRANSMEM 547 567 POTENTIAL.
FT TRANSMEM 583 603 POTENTIAL.
FT TRANSMEM 614 634 POTENTIAL.
FT TRANSMEM 678 698 POTENTIAL.
SQ SEQUENCE 751 AA; 84632 MW; A5C4F43540295EFC CRC64;

Query Match 13.4%; Score 387.5; DB 1; Length 751;
Best Local Similarity 25.8%; Pred. No. 6.4e-18;
Matches 106; Conservative 89; Mismatches 189; Indels 27; Gaps 9;

QY 127 TIVTEWNLVC-EDDWKAPLTISLFVGVLLGSFISGQSDRFGKRVLFVTMGQTGFSP 185
Db 303 SVQDFKMFCTKAYDAAWATVQIFGLGAIYGHLDHFGKRVSPFGISVGLFGV 362
QY 186 LQIFSKNEFMFVFLVGVGMQISNYAAAFVLGTEILGKSVRIIFSLGVGICFVAFGY-- 243
Db 363 ASGFAPSEVFAAFPIVGTSTIASILIVFYAYILEFTEPEQVRPLRS-----FENWGIAR 417
QY 244 MVLPLFAFYFIRDRMLLVALTPGVCVLMWFIPESPRWLSQGRPEAEVIRKAAKA 303
Db 418 LVFTLACFCYWRSAALATSLSLPLPVLLIPESPKWNTKRRFDARAAEKRVAWL 477
QY 304 NGIVVPTIFDPELOD-LSSKKQOOSHNLIDLLRTWIRWVMTIMLWMTISVGYFGLS 362
Db 478 SGIPYVNDQDSIESEKLEKSKTKYTKMDLFTSWTIAYRTIIVGVLSFSTLSAFGSD 537
QY 363 LDTPLNLDHIFVNCFLSAMVVPVAVLAWLLQVLP-----RRYSM--ATAFLGGSVLLF 416
Db 538 LNSGNLGNFVLSQVSGAVTAFKIFVFLDTPVDFDRRLHQQYQIAMILCYCVIMV 597
QY 417 MOLVPP-----DLYLTLVLMVGKFGVTAFAFMSVYVYTAELYPTVVRNMGVGVSS 467
Db 598 LMLPESDCGSGSDRLAI--IINIIGVSFEITWDACYLVAVECPFKTIRTIGIGTCS 655
QY 468 TASRLGSILSPFYVLGAYDFPLMGLS-LTILTAILLFLPESGTPL 517
Db 656 LIARTGALLAPQAMAYLSDIYPAPYAVVCSIGTISLLISCFLPDTRKGVDL 706

RESULT 5
YMP3_CAEEL STANDARD; PRT; 1222 AA.
AC Q10947;
DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL I39.9 KDA PROTEIN B0361.3 IN CHROMOSOME III.
GN B0361.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Du Z.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U00031; AAK18864.1; -.
DR WormPep; B0361.3; CE00752.
DR Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 51 71 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 232 252 POTENTIAL.
FT TRANSMEM 263 283 POTENTIAL.
FT TRANSMEM 307 327 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 355 375 POTENTIAL.
FT TRANSMEM 382 402 POTENTIAL.
FT TRANSMEM 413 433 POTENTIAL.
FT TRANSMEM 442 462 POTENTIAL.
SQ SEQUENCE 1222 AA; 139868 MW; CBA42A80A254FB80 CRC64;

Query Match 12.4%; Score 357; DB 1; Length 1222;
Best Local Similarity 21.1%; Pred. No. 9.6e-16;
Matches 118; Conservative 106; Mismatches 209; Indels 126; Gaps 14;

QY 3 DYDEVTAFIGEWPPORLIFFLSASIPNGFTGLSSVFLIA---TPHRCRVDPDAWL 58
Db 35 DPDKFVEAYGAYGKYQIFTYVVLQTL---NFFYSSSMYIMSFVQLNLEKQCE----- 83
QY 59 SSAMRNHTVPLRLDRDREVPKSCRRYRLATIANFSALGLEPGRDVLGQLEQESCLDGWE 118
Db 84 ---YKNETIP-----ISETCQ-----IETESSKAFGNLNGEYC-----G 114
QY 119 FSDQVVL-----STIVTEWNLVGEDDWKAPLTISLFVGVLLGSFISGQSDRFGK 170
Db 115 IAETLVNVTNQKASINLLVDFDLSCSHWFQFGLTIFTIGAVIAPFMSMLADRYGK 174
QY 171 NVLFVTMGQTGFSTQIFSKNFEMFVFLVGVGMQISNYAAAFVLGTEILGKSVRIIF 230
Db 175 PIIVTTAILAFLANMAASFPNFAIFLILRAFTGACSDSVLSVASVATCEYLSEKAR-AW 233
QY 231 STLGVCFIFAFGYVWLPLFAFYFIRDRMLLVALTMPGVLCVALWFWLPESPRWLSQGRF 290
Db 234 ITVVYNVAMSLGMVWTLTLVLTMTDDWRWRYFIVSLPGVGFALWFLPESPFWLITKNT 293
QY 291 EEAELVIRKAAKANGIWPSTIFDPSELQDLSKKQOOSHNLIDLLRTWIRWVMTIMSL 350
Db 294 EKLKVKYIKTNR----- 305
QY 351 WMTISVGYFGLSLDTPNLHGDIFVNCFL-SAMVEVPAYVLAWLQLQYLPVRYSM----- 403
Db 306 -MVISLVYFAISPMSEVELGSD-QVQAFLYSSLLIEIPAGLAVIPMMKMGKRMKVIWCLVF 363

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL; U29579; AAA69281.1; ALT_INIT.
 DR EMBL; AE000360; AAC75813.1; ALT_INIT.
 DR EcoGene; EGI3126; ygcS.
 DR InterPro; IPR003662; sub_transport.
 DR Pfam; PF00083; sugar_tr; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
 Complete proteome.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 57 77 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 115 135 POTENTIAL.
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 176 196 POTENTIAL.
 FT TRANSMEM 254 274 POTENTIAL.
 FT TRANSMEM 287 307 POTENTIAL.
 FT TRANSMEM 312 332 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 FT TRANSMEM 370 390 POTENTIAL.
 FT TRANSMEM 401 421 POTENTIAL.
 SQ SEQUENCE 445 AA; 48234 MW; B59E452721B15774 CRC64;

Query Match 10.78; Score 308; DB 1; Length 445;
 Best Local Similarity 26.68; Pred. No. 4.8e-13;
 Matches 102; Conservative 79; Mismatches 168; Indels 34; Gaps 11;
 QY 151 VGVLLGSGISQSDRFRGKKNVLFVTMGMTGFGFLOIFSKNEMFVFLVVGMOISN 210
 DB 65 LGLFGLSVLWISDHGRQKIFTFSLTLTLASFLQFFATTPHLLGLRLIGLGGD 124
 QY 211 YVAFVLGTEILGKSVRIIFSTLGV-CIFYAGVMVLPFA-YFIRO----WRMLLVALT 264
 DB 125 YSVGHTLLAEFSRRHGRIL-LGAFSVWTVGVVGLASIAGHHFISENPEAWRLASAA 182
 QY 265 MPGLCVLWVWFIPESRWLLSQGRFEAEVIRKAKANGIVVPSIFDPSEL--ODLS 322
 DB 183 LPALLITLLRWGTPESPRWLLRQGRFAEHAIVHR-----YFGPHVLLGDEVV 230
 QY 323 SKKQSHNILLTLRWIRMTVIMSIMLWMTISVGYFGLSDTLNHLGDIFVNCFLSAMV 382
 DB 231 TATHKHITLFSRYWR---RTAFNSVFFVCLVTPWFIYVITWLDPTIAQTIGLEDALTASL 287
 QY 383 EVPAYVLAWLLQYLPRIYSMATLFLGSGVLLF-----MOLVPPDLYLATVLMVVGKF 437
 DB 288 MLNALLIVGALLG-LVTLHLAHRKFLGLGSELLAATLVVMACLPSSGSLTLLLVLFST 346
 QY 438 GVTAFAFVWVYVTAELPTVVRNMGVSVSTASRLGSILSPYFV--VLGAYDRPLPYLM 495
 DB 347 TISAVSNLVGLTPAESFEPTDIRSLGVGFATMSRLGAAVSTGLLPWVLAOMGQVTLILL 406
 QY 496 GSLTILTAALT-LFLPESFGTLP 517
 DB 407 ATVLLGVFVITWLAPEIKALPL 429

RESULT 8
 YTI3_CAEEL STANDARD; PRT; 435 AA.
 AC Q10917;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE HYPOTHETICAL 48.6 KDA PROTEIN B0252.3 IN CHROMOSOME II.
 GN B0252.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae.
 CC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Du 2., Waterston R.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 DR EMBL; U23453; AAC46757.1; .
 DR WormPep; B0252.3; CE02419.
 DR InterPro; IPR003662; sub_transport.
 DR Pfam; PF00083; sugar_tr; 1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 55 75 POTENTIAL.
 FT TRANSMEM 384 404 POTENTIAL.
 SQ SEQUENCE 435 AA; 48571 MW; 5F6160359FA1840B CRC64;

Query Match 10.18; Score 291.5; DB 1; Length 435;
 Best Local Similarity 21.5%; Pred. No. 5.4e-12;
 Matches 90; Conservative 78; Mismatches 168; Indels 83; Gaps 8;
 QY 128 IVTEWNLVCEDDWKAPLITISLFFVGVLLGSFTSGQLSDRFGKKNVLFVTMGOTGFSFLQ 187
 DB 91 VADEFDLTGDAWLAESTITTFYVGVNMGIMFTPLADHYGRLPVFVAIVLLMAYGGMIS 150
 QY 188 IFSKNEMFVFLVVGMOISNYAAFLVGLTEILGKSVRIIFSTLGVCFIFYAGVMVLP 247
 DB 151 AFSTIMMFCIMRMIRGIFYTAAGLWGLVGYENTPLRLR-EFTSVYFGVMWVGACFLG 209
 QY 248 LFAYFIRDMRLVALTMPGV-LCVALWVWFIPESRWLLSQGRFEAEVIRKAKANGI 306
 DB 210 LLAAYLPDWRVLMFCISVNFVALLIYMTVPESLHFLVSSQOONEKIEAWLEKIRPKGD 269
 QY 307 VPESTIFDSELDQLSKKQSHNILLTLRWIRMTVIMSIMLWMTISVGYFGLSLDTP 366
 DB 270 ISASDI-----VEDRDENGSSFKTLCRE----- 292
 QY 367 NLHGDIFVNCFLSAMVEVPAYV-LAWLLQYLPRIYSMATLFLGSGVLLFMQLVPPDLY 425
 DB 293 -----IKTFKTLTQDRRYTYLFWI-----GILLYEFGKPLFE 326
 QY 426 YLATVLMVVGKGVTAFAFVWVYVTAELPTVVRNMGVSVSTASRLGSILSPYFVYILGA 485
 DB 327 FCA-----HEFGRSSLSHFFSDFHEQIFPTDGRNKCIGFCETLSRFGGMLSPYLSHLTA 380
 QY 486 YDRFLPYLLMGSLTILTAITLTLFLPESFGTLPDITDQMLRVKGMKHKRTPSHTRMLKD 544
 DB 381 VHAPALPILSLIAVSGGLTLLLPETLNKLPSTI-----AETASRRQLIDD 428

RESULT 9
 YCEI_BACSU STANDARD; PRT; 400 AA.
 AC Q34591;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YCEI.
 GN YCEI.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]

```
RP SEQUENCE FROM N.A.
RC STRAIN=166;
RA Kumano M., Tamakoshi A., Yamane K.;
RT "A 32 kb nucleotide sequence from the region of the lincomycin-
RT resistance gene (22-25 degree) of the Bacillus subtilis chromosome and
RT identification of the site of the lin-2 mutation.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB000617; BAA22256.1; -
CC DR EMBL; 299105; CAB12089.1; -
CC DR Subtilist; BG12773; ycel.
CC DR InterPro; IPR003662; sub.transprtr.
CC DR Pfam; PF00083; sugar.tr.1.
CC DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
CC DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
CC FT TRANSMEM 15 35 POTENTIAL.
CC FT TRANSMEM 50 70 POTENTIAL.
CC FT TRANSMEM 78 98 POTENTIAL.
CC FT TRANSMEM 99 119 POTENTIAL.
CC FT TRANSMEM 143 163 POTENTIAL.
CC FT TRANSMEM 166 186 POTENTIAL.
CC FT TRANSMEM 218 238 POTENTIAL.
CC FT TRANSMEM 254 274 POTENTIAL.
CC FT TRANSMEM 281 301 POTENTIAL.
CC FT TRANSMEM 305 325 POTENTIAL.
CC FT TRANSMEM 344 364 POTENTIAL.
CC FT TRANSMEM 371 391 POTENTIAL.
CC SQ SEQUENCE 400 AA; 43708 MW; EOAE0CEB5DD27395 CRC64;

Query Match 10.0%; Score 287; DB 1; Length 400;
Best Local Similarity 24.9%; Pred. No. 9.7e-12;
Matches 104; Conservative 76; Mismatches 161; Indels 76; Gaps 18;

QY 116 GNEF-SQDV-YLSTIV----TEWNLVCD-DWKAPLTISLFVGVLLGSFTSGOLSDRF 168
DB 19 GWFDAVDGILSFIIAALHVENNLSPPEMKWG-----SVNSIGMAGAFLEGLADRI 74
QY 169 RKNVLFTVMGQTGSPFLQIFSKNFEMFVLVGVGMQISNYVAAFVLGTEIL-----G 223
DB 75 RKKVFITILLCFSIGSGISAFVTSLSAFLIRFVIGMGLGELPVASTLVSEAVVPEKRG 134
QY 224 KSVRIEFTGLVCIFYAGYVWVLPFAVFI---RDWRMLVALTMPGVLCVALWFFIPES 280
DB 135 RVIVLES-----FWVGWLAALISVIFPSFGWQAALLLTALTAFYALYLRSLPDS 188
QY 281 PRLWISQGRFEAEVIRKAANGIVVPSTIFDPSELQDLSKKQOS--HNILDLLRTWN 339
DB 189 PKY-----ESLSAKKRSWENVKSWARQY 213
QY 340 IRMVTIMSLMWTLSVGYFGLSLDTPN---LHGDIFVNCF----LSAMVEVPYVLAWL 392
DB 214 IRPTVMSI-VNFCVVFYSYGNLWLPSPVMLLKGFMSIQSFYVLLMTLAOLPGYFSAW 272
QY 393 LLOYLPRYSNATALF-LGGSVLLFMQLVPPDLVYLATVLMVGKFGVTAFAFSVVYTA 451
DB 273 LIEKAGRWILVVLIGTAGSAFYFG--TADSLSLLLTAGVLLSFFNL-GAWGVLYATP 329
QY 452 ELYPTVRNMVGVSSTASRLGSILSPYFV-YLGAYDRFLPYILMGSITILTAILTL 507
DB 330 EQYPTPAIRATSGTFAFGRIIGGIFGLLVGTAA--RHISFSVIESFICITAILAV 384
```

```
RESULT 10
YAAU_ECOLI STANDARD; PRT; 443 AA.
AC P31579; P31578; P75628;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YAAU.
GN YAAU OR B0045.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS
CC IN POSITION 142 AND 232 THAT PRODUCE TWO SEPARATE ORFS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D10483; BAA01321.1; ALT_FRAME.
CC DR EMBL; AE000114; AAC73156.1; -
CC DR PIR; S40566; S40566.
CC DR EcoGene; EG11566; yaaU.
CC DR InterPro; IPR003662; sub.transprtr.
CC DR Pfam; PF00083; sugar.tr.1.
CC DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
CC DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
CC KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
CC Complete proteome.
CC FT TRANSMEM 19 39 POTENTIAL.
CC FT TRANSMEM 54 74 POTENTIAL.
CC FT TRANSMEM 85 105 POTENTIAL.
CC FT TRANSMEM 114 134 POTENTIAL.
CC FT TRANSMEM 146 166 POTENTIAL.
CC FT TRANSMEM 174 194 POTENTIAL.
CC FT TRANSMEM 242 262 POTENTIAL.
CC FT TRANSMEM 283 303 POTENTIAL.
CC FT TRANSMEM 323 343 POTENTIAL.
CC FT TRANSMEM 374 394 POTENTIAL.
CC FT TRANSMEM 402 422 POTENTIAL.
CC FT CONFLICT 7 7 F -> L (IN REF. 1).
CC FT CONFLICT 12 16 FSSIH -> CLHY (IN REF. 1).
CC FT CONFLICT 30 30 D -> Y (IN REF. 1).
CC FT CONFLICT 115 115 L -> F (IN REF. 1).
CC FT CONFLICT 132 132 T -> N (IN REF. 1).
```

```

FT CONFLICT 240 240 R -> C (IN REF. 1).
FT CONFLICT 304 304 N -> S (IN REF. 1).
FT CONFLICT 395 395 N -> D (IN REF. 1).
SQ SEQUENCE 443 AA; 48666 MW; AEF4848B109DE05 CRC64;

Query Match 9.3%; Score 267; DB 1; Length 443;
Best Local Similarity 24.6%; Pred. No. 2.1e-10;
Matches 104; Conservative 81; Mismatches 184; Indels 54; Gaps 13;

QY 134 LVCEDDKAPLATISLFVGVLLGSFIS-----GQSDRFGRKKNVLFVTMGMTGFSFLQIF 189
Db 48 LKLDADM-----IGLLGAGTLAGLFVGTSLFQISDKVGRKMFLLIDIIAIGVISVATMF 102

QY 190 SKNFEMFVFLVGVLLGQISNYAAVFLVGTETLLGKSVRIIFSTGLVCIFYAFGYVLPFLF 249
Db 103 VSSPVELLMVRVLIGVIGADYPIATSMITFSSTRQR-AFSISFIAAMVYVGCATCADLV 161

QY 250 AYFRD-----WRMLVALTMPGVLCVALWFIPESPRWLISQGRPEEAETIRKKAANG 305
Db 162 GYWDYVEGGWRMGLGSAIPCLLIILIGREFELPESRWLLRGRVKECEEMIK----- 215

QY 306 IVPSTIFDPSEIQ-----DLSSKKQOQSHNLDLLRTWIRVMTIMSLMWTISVGYFG 360
Db 216 LFCEPVAFDEEQQOTFRDLNRRHPEF-VLFVAAIWTQVIMPEAIYTFGQIVGLLG 274

QY 361 LSLDTPNLHGDIQVNFCLSAMVEPAVLANLLQLVLPYRYMWATA-----LFLGGSVLLF 416
Db 275 LGVGKNAALGNVVISLF-----FMLGCIPTPMLNLTAGRRPLLGSPAMMT 320

QY 417 MOLVPPDLYLATVLMVGVKFGVTAIFS-----MVVYVYTAELYPTVVRNMGVSVSTASR 471
Db 321 LALAVALGLIPDMGIWLVVMAFAVAFSGGPGNLQWLINELPPTDIRASAVIMSLSR 380

QY 472 LGSILSPYF--VVLGAYDRELPYLMGSLTILTAIL--TLFLPESFGTPLPDTIDQMLRV 527
Db 381 IGTIVTWPALPIFINNYG-ISTNMLMGAGISLFGLLISVAFAPETRGMSLAQISNMTIRG 439

QY 528 KGM 530
Db 440 QRM 442

RESULT 11
YDJK_ECOLI STANDARD; PRT; 459 AA.
AC P76230; P76911;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YDJK.
GN YDJK OR B1775.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,

```

```

RA Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.,
RT "A 570-kb DNA Sequence of the Escherichia coli K-12 Genome
DNA Res. 3:363-377(1996).
CC Corresponding to the 28.0-40.1 min Region on the Linkage Map.";
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000272; AAC74845.1; .
DR EMBL; D90821; BAA15573.1; ALT_INIT.
DR EcoGene; EGI3487; ydjk.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
FT TRANSMEM 113 133 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 272 292 POTENTIAL.
FT TRANSMEM 302 322 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 421 441 POTENTIAL.
SQ SEQUENCE 459 AA; 49602 MW; BC8AB53ECB8DD77 CRC64;

Query Match 9.2%; Score 265; DB 1; Length 459;
Best Local Similarity 26.2%; Pred. No. 2.9e-10;
Matches 101; Conservative 79; Mismatches 153; Indels 52; Gaps 18;

QY 152 GVLGSGTSGQLSDRFGRKKNVLFVTMGMTGFSFLQIFSKNFEMFVFLVGVLLGQISNY 211
Db 70 GMFLGALVGGIIGDKTGRNFIYEAHIASMVYVGAESPNDFLIACRFVNGVGLGALL 129

QY 212 VAAFLVLTETIL-GK-----SVRIIFS---TLGVCIFYAFGYVLPFLFAFYTRDWRMLVA 262
Db 130 VTLFAGTYMFGNRGTWSSRVFIGNWSYPLCSLIANG--LTPLIS-AEWNVRVQLLI 186

QY 263 LTPMGVLCVAL-WFIPESPRWLSIQGRPEEAETIRK-----AAKANGIVVPTIFDPS 317
Db 187 PALSIALATLAWRYFPESPRWLESRGVYQAEKVMRSIEEGVIRQTGKPLP-----PW 241

QY 318 LODLSKKQOQSHNLDLLRTWIRVMTIMSLMWTISVGYFGSLDTP-----NLH 369
Db 242 IAD-DGKAPQAPVPSALLTGVLKRVILGSCVL-TAMNVVQYTLINWLPTIFMTGQINLK 299

QY 370 GDFVNCFLSAMVEVP-AVLAWLLOLPLRPYRSMATALLFGSVLLFMQLVPPDLYYLA 428
Db 300 DSIVLNTM--SMFGAPFGFIAMLVMDKIPRK-TMGVGLLIILAVLG-----IYSLQ 349

QY 429 T---VLVMVGKFGVTAAFSNV-----VYVYTAELYPTVVRNMGVSVSTASRLGSLSPY--- 479
Db 350 TSMLLITLIGFFLITFVYVYVVCYASAVYVPEIWPTEAKLRGSGLANAVGRISGIAAPYAV 409

QY 480 FVYLGAIDRELPYLMGSLTILTAI 504
Db 410 AVLLSSYGVGTGVFILLGAVSIIVAI 434

```

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D50543; BAA09111.1; -
 DR EMBL; Z99108; CABI2635.1; -
 DR Subtilist; BG11854; yfiG.
 DR InterPro; IPR003663; Sugar_transportr.
 DR InterPro; IPR003662; sub.transports.
 DR Pfam; PF00083; sugar.tr.1.
 DR PRINTS; PR00171; SUGTRNSPORT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 30 50
 FT TRANSMEM 60 80 POTENTIAL.
 FT TRANSMEM 93 113 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 156 176 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 FT TRANSMEM 302 322 POTENTIAL.
 FT TRANSMEM 332 352 POTENTIAL.
 FT TRANSMEM 353 373 POTENTIAL.
 FT TRANSMEM 401 421 POTENTIAL.
 FT TRANSMEM 424 444 POTENTIAL.
 FT SEQUENCE 482 AA; 52756 MW; 90851C4F4C48BE01 CRC64;

Query Match 8.9%; Score 256; DB 1; Length 482;
 Best Local Similarity 23.9%; Pred. No. 1.2e-09;
 Matches 105; Conservative 83; Mismatches 167; Indels 84; Gaps 19;
 QY 144 ITISLFVGVLLGSFSTISGQLSDFGRKN-----VLFF--TMGMQGFSLQIFSKNFEM 195
 Db LVASSLLGLAFAFGMFGRLSDRGRRKTKLYLALLFIAATLGC-----TFSPNASV 115
 QY 196 FVFLFVLGM--GQISNVAAVFLGTEILGKSVRII-----FSTLGVCFYAFGYMWLPL 248
 Db 116 MIAPFRLGLAVGCASVTPTFLAETSPAERRGRIVTQNELMIVIGQLLAYTFNALIGST 175
 QY 249 FAYFIRDMRLVATLTPGVLCVALLW---FIPSPRLISQGRFEAEVIRKAANKANG 305
 Db 176 MGESANVRWMLVATLTPAVV---LWFGMLIVPESPRLAAGRMGDALRVLRQIRE--- 229
 QY 306 IVPSTIFDPSELQD-----SSKKQOSHNLDLRTNIRMTI---MSIMLWMT 353
 Db 230 -----DSQAQOEIKHAKTEGAKKAGFH---DFQEPWIRRLIFIGIGIAIVQOIT 278
 QY 354 --ISGVFGSL-----DTPNLHGDFVNCFLSAMVEVPAYVLAWLILQYLPRR--- 400
 Db 279 GVNSIMTYGTEILREAGQFTEAALIGNI-----ANGVISVIAVIFGILWLGKVRPRMLI 333
 QY 401 ---YSMATFLGSLVLEWLPDLYLATVLMVKGFGVTAARSMVYVYTAELYPTV 457
 Db 334 IGQIGTMTALLIGILSILEGTPALYVVLSTILFLAQQTATISVITWMLSEIFPMH 393
 QY 458 VRNMGVGVSS-----TASRLSILSPYFV-YLGAYDRFLPYLMGSLTILFAITLFLPES 512
 Db 394 VRGLGMGISTCLMTANFLGTFTEPILLNHGMSATFFIEVAMNAIL--FVKVYVET 451
 QY 513 FGTPLPTIDQMLRVKGMK 531
 Db 452 KGRSL-EQLEHSFRQYGR 469
 PRT; 452 AA.
 STANDARD;
 ID YDJE_ECOLI
 AC P38055; P77244;

RESULT 14
 YDJE_ECOLI
 ID YDJE_ECOLI
 AC P38055; P77244;

DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL METABOLITE TRANSPORT PROTEIN YDJE.
 CC YDJE OR B1769.
 OS Escherichia coli.
 CN Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
 Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
 Sampei G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horluchi T.;
 RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 RN [3]
 RP SEQUENCE OF 360-452 FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=89357501; PubMed=2670682;
 RA Jerlstroem P.G., Bezjak D.A., Jennings M.P., Beacham I.R.;
 RT "Structure and expression in Escherichia coli K-12 of the
 L-asparaginase I-encoding ansA gene and its flanking regions.";
 RL Gene 78:37-46(1989).
 RN [4]
 RP IDENTIFICATION.
 RX MEDLINE=95075659; PubMed=7984428;
 RA Borodovsky M., Rudd K.E., Koonin E.V.;
 RT "Intrinsic and extrinsic approaches for detecting genes in a
 bacterial genome.";
 RL Nucleic Acids Res. 22:4756-4767(1994).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE000272; AAC74839.1; -
 DR EMBL; D90820; BAA15560.1; -
 DR EMBL; D90821; BAA15567.1; -
 DR EMBL; M26934; -; NOT_ANNOTATED_CDS.
 DR EcoGene; EG12369; ydje.
 DR InterPro; IPR003662; sub.transports.
 DR Pfam; PF00083; sugar.tr.1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
 KW Hypothetical protein; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 60 80 POTENTIAL.
 FT TRANSMEM 96 116 POTENTIAL.

FT	TRANSMEM	119	139	POTENTIAL.
FT	TRANSMEM	167	187	POTENTIAL.
FT	TRANSMEM	264	284	POTENTIAL.
FT	TRANSMEM	299	319	POTENTIAL.
FT	TRANSMEM	327	347	POTENTIAL.
FT	TRANSMEM	350	370	POTENTIAL.
FT	TRANSMEM	397	417	POTENTIAL.
FT	TRANSMEM	418	438	POTENTIAL.
SEQ	SEQUENCE	452 AA;	50318 MW;	5A596DEA229205C2 CRC64;

Query Match	8.9%;	Score 255.5;	DB 1;	Length 452;
Best Local Similarity	24.0%;	Pred. No. 1.2e-09;		
Matches 114;	Conservative 92;	Mismatches 160;	Indels 109;	Gaps 24;

Qy	84	YRLATIANFSGALGLEPGRDVLGQLEQESCILDGW-EFSQDYYVLSTIVTE-WNLVCEDDWK	141
Db	21	YRIFCIISFLL-----LTGFLSYSGNVVLAKLVNGWS-----NNFLN	59
Qy	142	APLTISLFEVGVVLGSGFTSGOLSDRGRK-----NVLEVTMGMTGTFSLQIFSKNFEMF	196
Db	60	AAFTSALMF-GYFISGLTGGTIGDYFGRRAFRNLLIVGIA-ATGAAFVP-----DMY	111
Qy	197	VVLF-----VLVGMQISNYVAAFVLGTEILGKSVRIIF-STLGVCIPIAFYQ	243
Db	112	WLIFPRFLMGTMGALIMVGYASFEFIPAVRGR-----WSARLSFVGNWSPMLSAALCV	167
Qy	244	MVLPLFAFYIRDRMLLVALTMPGVLCAVMWFIP-----ESPRWLISQGRFEAEVIR	298
Db	168	VVIAPFS-----WRIMFLL-----GGIGILLANFLSGKFIESPRWLACKQGIAGAEQOLR	218
Qy	299	KA-----AKANGIVPSTTFDPSLEQDLSKKQKOSHILDLRTNIRMVTIMSLMWMTI	354
Db	219	EVEQOIEREKSIRLP-----PLTSYQSNSKKVKIKGTWELFKGEMLRRTLVIAITVLIAM	273
Qy	355	SVGVFGLSDPNLHGDIENV-----CFLSAMVEVPAYV---LAWLLLOYLPRRYSM	403
Db	274	NISLYTITWIPT-----IFVNSGIDVDPKSLIMTAVIMIGAPVGIFIAALIIDHFPFR---	326
Qy	404	ATALFLGGSVLLFQVLPPDLYLATVLVMVGKFGVTAAFSM-VV-----VYTAELPYPT	456
Db	327	-----LFGSTELLIIAVLGYIYSIQTTETWALLIYGLVMVIFLYMVVCFASAVIPELWPT	381
Qy	457	VVRNMGVGSVTSASRLSGILSPYFV---YLGAYDRFLPYILMGSTIILTA-ILTLF	508
Db	382	HLRUGSGFVNAVGRIVAVFTFPGVGAALLTHYGSITVFMVLGVMLGMLCALVLSIF	436

RESULT	15	
PCAK_ACICA		
ID	PCAK_ACICA	STANDARD;
AC	AC43975;	PRT; 457 AA.
DT	15-JUL-1999 (Rel. 38, Created)	
DT	15-JUL-1999 (Rel. 38, Last sequence update)	
DT	20-AUG-2001 (Rel. 40, Last annotation update)	
DE	4-HYDROXYBENZOATE TRANSPORTER.	
DE	PCAK.	
GN	Acinetobacter calcoaceticus.	
OS	Acinetobacter calcoaceticus.	
OC	Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;	
OC	Acinetobacter.	
OX	NCBI_TaxID=471;	
[1]	RN	
RP	SEQUENCE FROM N.A.	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BD413 / ADP1;	
RC	MEDLINE=94341565; PubMed=8063101;	
RA	Kowalichuk G.A., Hartnett G.B., Benson A., Houghton J.E., Ngai K.-L.,	
RA	Ornston L.N.;	
RT	"Contrasting patterns of evolutionary divergence within the	
RT	Acinetobacter calcoaceticus pca operon.";	
RL	Gene 146:23-30(1994)	
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE	
CC	(POTENTIAL).	
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.	

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L05770; AAC37151.1; -.
DR InterPro; IPR003662; sub_transprot.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transport; Transmembrane; Inner membrane.
FT TRANSMEM 35 55 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 169 189 POTENTIAL.
FT TRANSMEM 192 212 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 311 331 POTENTIAL.
FT TRANSMEM 339 359 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT TRANSMEM 401 421 POTENTIAL.
FT TRANSMEM 427 447 POTENTIAL.
SQ SEQUENCE 457 AA; 49277 MW; 4F585FF77361A1567 CRC64;

```

```

Query Match      8.8%; Score 252.5; DB 1; Length 457;
Best Local Similarity 24.9%; Pred. No. 1.1e-09;
Matches 113; Conservative 78; Mismatches 177; Indels 85; Gaps 20;

Qy 110 QESCLDGWEFSQDYVLGFTIVTEW--NLVC-----EDDW---KAP 143
Db 15 QRSLSDAQALINDAPLSRY--QWLLAIVCFLIVFDGIDTAAMGFTAPALAAQMGVDRSQ 72

Qy 144 L--TISLFFVVLGTSFISGOLSDRFGKKNVLFVTMGMTQGFSLQIFSKNFEMFVFLV 201
Db 73 LGPVMSAALGMIITGALVSGPTADRFGRKIVLSMSLVFGGFTLACAYSTNLSLVIFR 132

Qy 202 LVGMGQTSNVVAAPVLGTEILIGKSVRIIFTGLGVCIFYAFGYVMVLPPLFAFYFIDW---- 256
Db 133 LTGIGLGAAMPNATTLSEYCPARIR---SLLYTCMP--CGYNLGNWAGGFISSWLIAP 187

Qy 257 ---RMLVALVTMPGVLCVALWVFIPESRWLIISGR--FEAEAVIIRKAA--KANGIVVPS 310
Db 188 GWHSLFLTGGWAPLIMLLVIFFLPESYRELIVKGNKTKVROILSRIPAKQVQGV---T 244

Qy 311 TIFDPSLEQLDSSKK-----QQSHNTLDLRTWNIRWVTIMSLMWTMTISVGVFGL 361
Db 245 EFHVPPEEKVEAGTKKGVFGLFSKYYKGVTVLWVYFVGLVMYIYLLTWSLPTLMRETGA 304

Qy 362 SLDTPNLHGDI F-VNCFPLSAMVEVPAYVLAWLLLOYLPRR-----YSMATALFICGSVLLF 416
Db 305 SLERAFLGLGLFGGVVISAL-----FIGWANDRENPNRIIAGFYLAAGIF---AVIVG 355

Qy 417 MQLVPPDLYLATVLMVMVGFGVTAAFSVMVYVYTAELYPTTVVNMGMGVSYSTASRLGSIL 476
Db 356 QLSLNPTL--LALFILCAG--IAVNGAQSSMPVLARSFYTQCRATGVAMWSGIGRFCAV-- 411

Qy 477 SPFYVILGAYDRELPLVILMGSLTILTAITLFL 509
Db 412 --FGAWIGA-----VLLGNWWSFTMILSMILI 435

```

Search completed: January 25, 2002, 10:15:21
Job time: 327 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2002, 10:14:54 ; Search time 47.78 Seconds
(without alignments)
1705.184 Million cell updates/sec

Title: US-09-521-195-3
Perfect score: 2883
Sequence: 1 MRDYDEVTAFLGEGWGPORL.....HTRMLKDGQERPTILKSTAF 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL17:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2295.5	79.6	564	11 Q9WTN6	Q9wt6 mus musculus
2	2227	77.2	551	4 Q9H015	Q9h015 homo sapien
3	2218	76.9	551	4 O14546	O14546 homo sapien
4	2157	74.8	553	11 Q9R141	Q9r141 rattus norv
5	2148	74.5	553	11 Q9Z306	Q9z306 mus musculus
6	885	30.7	548	5 Q9VCA2	Q9vca2 drosophila
7	881	30.6	548	5 O01384	O01384 drosophila
8	816	28.3	567	5 Q9VCA3	Q9vca3 drosophila
9	777	27.0	568	5 Q9U539	Q9u539 caenorhabdi
10	777	27.0	576	5 O02270	O02270 caenorhabdi
11	753.5	26.1	561	5 Q9V6L6	Q9v6l6 drosophila
12	740	25.7	554	4 Q9NQD4	Q9ngd4 homo sapien
13	738.5	25.6	556	4 O75751	O75751 homo sapien
14	738	25.6	554	4 O15395	O15395 homo sapien
15	734	25.5	554	4 O15245	O15245 homo sapien
16	733	25.4	593	11 P70485	P70485 rattus norv
17	728	25.3	593	11 P97558	P97558 rattus norv
18	727	25.2	555	4 O15244	O15244 homo sapien
19	724	25.1	555	11 Q9R0W2	Q9r0w2 rattus norv

20	723.5	25.1	554	6 O77504	O77504 oryctolagus
21	718	24.9	556	11 Q63089	Q63089 rattus norv
22	716.5	24.9	674	5 Q9VIK2	Q9vik2 drosophila
23	714.5	24.8	553	11 O70577	O70577 mus musculus
24	711	24.7	554	6 O02713	O02713 sus scrofa
25	710	24.6	535	11 O63314	O63314 rattus norv
26	709	24.6	556	11 O08966	O08966 mus musculus
27	704	24.4	556	11 Q9RIQ4	Q9riq4 mus musculus
28	703.5	24.4	539	4 Q9H2W5	Q9h2w5 homo sapien
29	703	24.4	551	4 Q9Y226	Q9y226 homo sapien
30	690.5	24.0	551	11 Q9WTW5	Q9wtw5 mus musculus
31	690	23.9	548	4 Q9Y694	Q9y694 homo sapien
32	688.5	23.9	551	11 Q99JF0	Q99jf0 mus musculus
33	681.5	23.6	551	11 O88446	O88446 rattus norv
34	675.5	23.4	562	13 O57379	O57379 pseudopleur
35	660.5	22.9	456	4 O14567	O14567 homo sapien
36	660.5	22.9	537	11 O88909	O88909 mus musculus
37	656.5	22.8	536	11 Q9RIU7	Q9riu7 rattus norv
38	651	22.6	557	5 Q9VEX8	Q9vex8 drosophila
39	650	22.5	538	5 Q9V6H5	Q9v6h5 drosophila
40	643.5	22.3	551	11 O35956	O35956 rattus norv
41	638.5	22.1	762	5 Q3VNW8	Q3vnx8 drosophila
42	638	22.1	545	11 Q61185	Q61185 mus musculus
43	628	21.8	563	4 O95742	O95742 homo sapien
44	627	21.7	575	5 Q9BKR9	Q9bkr9 caenorhabdi
45	621	21.5	569	5 Q9VEY2	Q9vey2 drosophila

ALIGNMENTS

RESULT	1				
Q9WTN6					
ID	Q9WTN6	PRELIMINARY;	PRT;	564	AA.
AC	Q9WTN6;				
DT	01-NOV-1999 (Tremblrel. 12, Created)				
DT	01-NOV-1999 (Tremblrel. 12, Last sequence update)				
DT	01-JUN-2001 (Tremblrel. 17, Last annotation update)				
DE	OCTN3.				
GN	SLC22A9 OR OCTN3.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=TESTIS;				
RA	Nezu J.;				
RT	"Mouse OCTN3 - a novel OCTN transporter family protein.";				
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.				
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).				
CC	-!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.				
DR	EMBL; AB018436; BAA78343.1; -.				
DR	MGD; MGI:1929481; SLC22a9.				
DR	InterPro; IPR003662; sub.transporter.				
DR	Pfam; PF00083; sugar.tr.1.				
KW	PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.				
SQ	Transmembrane.				
	SEQUENCE 564 AA: 63320 MW: C37FDCA6395DAD01 CRC64;				

Query Match 79.6%; Score 2295.5; DB 11; Length 564;
Best Local Similarity 80.6%; Pred No. 7.5e-136;
Matches 435; Conservative 54; Mismatches 48; Indels 3; Gaps 1;

QY	1	MRDYDEVTAFLGEGWGPORLIFLLSASIPNGFTGLSVFLIATPEHRCRPVDAANLSS	60
Db	1	MLDYDEVTAFLGEGWGPORLIFLLSASIPNGFTGLSAVFLTAPEHRCRPDPVNLSS	60
QY	61	AWRNHTVPLRLKDGREVPKCRRYRLATIANFSLGLPGRDVDLGOLEQESCLDGWEFS	120
Db	61	AWRNHSIPMTKDGPEVPKCRRYRLATIANFSELGLEPGRDVDLEQLEQENCLDGWEYD	120

Qy	121	ODVYLTSTIVTEWNLVCEDDKKAPLITSLTFFVGVLGGSFISGQLSDRGFRGNLVFTVMGMQ	180
Dd	121	::: ::: : :: : :: : :: : :: : :: :	
Db	121	KDIFELSTIVTEWDLVKCKDDKNKAPLTTSFYVGVLLGSFISGQLSDRFGRKNTILFLTMAMH	180
Qy	181	TGFSFIQIFSKEPFMFVVLFVLCMGQISNYAAAFVLGTGLEICKSVRIIIFSLGVCIFYA	240
Dd	181	:: : :: : :: : :: : :: : :: : :: : :: :	
Db	181	TGFSFIQVFSENVFEFPLLTYTLVCMGHSHINYAAAFVLGTLEMLSKSVRIIIFATLGVCIFFA	240
Qy	241	FGYMVLPLFAFYFDWRMLLVALTMPGLCVLAWMWFIPSPRWLISSQGFEAEVIIRKA	300
Dd	241	:: : :: : :: : :: : :: : :: : :: : :: :	
Qy	301	AKANGIYVPSTIDPSE---LDLSKKQQOSHNIIDLRLTNIMVTIMSILMMWTISVG	357
Dd	301	:: : :: : :: : :: : :: : :: : :: : :: :	
Qy	358	YFGLSLDTPNLHGDI FVNCFLSAWVEPAVVAWLALLQVLPREYSNATALEFGSGSVLLFM	417
Dd	361	YFGLSLDTPNLNGNIYNVNCPELLAAVEPAVVAWLALLQHVSRRYSAGSULFLGGSSVLLV	420
Qy	418	OLVPPDLLYLATLVMYGKGFTAAFESMVVYYTAELYPTVVRNMVGVSASTASRLGSI LS	477
Dd	421	OLVPSDLHYLSTLLVMYGKFGITSAYSMVVYYTAELYPTVVRNMVGVSASTASRLGSI LS	480
Qy	478	PYFVYLCAYDRFLPYILMSGSLTILTALILFLPESTGTPDPOTIDOMLRVKGMKHRKTPS	537
Dd	481	PYFVYLCAYDRRLPYILMSGSLTILTAITLTFPPESSGVSLPETIDEMQKVKLQRQSLS	540
RESULT	2		
Q9H015	ID	PRELIMINARY; PRT; 551 AA.	
AC	Q9H015;		
DC	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)		
DE	INTEGRAL MEMBRANE TRANSPORT PROTEIN.		
OS	UT2H.		
GN	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
RX	NCBI_TaxID=9606;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RA	Spritzberger F., Gruendemann D., Schoemig E.;		
RL	Submitted (DEC-1996) to the EMBL/GenBank/DDJB databases.		
CC	- - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	- - SIMILARITY: TO SUGAR TRANSPORTER FAMILY.		
DR	ENBL; Y09881; CAA71007.1; -		
DR	InterPro; IPR003662; sub_transporter.		
DR	Pfam; PF00083; sugar_tr; 1.		
SW	SEQUENCE 551 AA; G2155 MW; C827A99AA78C9443 CRC64;		

Query Match	77.28;	Score 2227;	DB 4;	Length 551;
Best Local Similarity	76.3%;	Pred. No. 1.4e-131;		
Matches 425;	Conservative 57;	Mismatches 69;	Indels 6;	Gaps 3;
QY 1	MRDYDEVTAFLGEWGPFORLFIFFLLISASITPNGSTGLSSVFLIATPEHRCRVPDAANLSS	60		
Db 1	MRDYDEVTAFLGEWGPFORLFIFFLLISASITPNGSNGMVFLAGTPEHRCRVPDAANLSS	60		
QY 61	AWRNHTVPLRLDRGREVPHSCRRYRLATIANFSALGLEPCRGVDLGLQEESCLDGEWFS	120		
Db 61	AWRNNSVPLRLDRGREVPHSCRRYRLATIANFSALGLEPCRGVDLGLQEESCLDGEWFS	120		
QY 121	QDYLSTIVTEWNLYCEDDWKAPITISLTFVGVLLGSFISQGLSDRFGKKNLVFTMGMQ	180		
Db 121	QDYLSTIVTEWNLYCEDNWKVPITSLTFVGVLLGSFISQGLSDRFGKKNLVFATMAVQ	180		
QY 181	TGFSFLQIFSKNFEHMFVFLVFLVGMGQISNVAAFVLGTEILGKSVRIIFSTLGVCIETFA	240		

Db	181	TGFSFLQIFSISWEMFTVLVFI	VGMGQISNVVVAFLTGLTGILGKSVRIIFSTLGVCTFFA	240
Qy	241	FGYMWLPFAFYFIDRMRLLVAT	MPGVLCAVLMWFIPESPRWLISQGRFEAEVIIRKA	300
Db	241	VGIMLLPLFAFYFIDRMRLLVAT	MPGVLCAVLMWFIPESPRWLISQGRFEAEVIIRKA	300
Qy	301	AKANGIVVPSTIFDPSELQDSSK	QOQSHNLDLLRTWNIWMYTIMSLWMTISVGYFG	360
Db	301	AKMNNIAVPAVIED--SVEELN	PLKQOKAPILDLFRTNIAIMTISLLIWMILTSVGYFA	358
Qy	361	LSLDPNLHGHDIFVNCFLSAMVE	PAPVAVLAWLLQVLPYPRYSNATATFLGGSVLLFMQLV	420
Db	359	LSLDAPNLHGDAYLNCFLSALIE	IPAYITANLLRLTPRYIIAANVLFWGGGVLLFIQLV	418
Qy	421	PPDIYLIATVLVMVGKFGVTAAS	VMVYTAELYPTVVRNMGVGVSTASRLGSLSPYF	480
Db	419	PVDYFYISLIGVLMGKFGITS	SAFSMLVYFTAELYPTLVNRNMAVGVTASRVGSIITAPYF	478
Qy	481	VYLGAYDRFLPYILMGSLTITL	TATLTFLPESFGTLPDPTIDOMLRVGMKHKRKTFSHTR	540
Db	479	VYLGAYNRMLPYIVMGSLSLV	LIGLITLFFPESLGMILPETLEQMKVKNFRSCK---	KTR 535
Qy	541	MLKDGQRPITLKSTAF	557	
Db	536	DSMETEENPKVL-ITAF	551	
RESULT	3			
ID14546		PRELIMINARY;	PRT;	551 AA.
AC	014546			
DT	01-JAN-1998	(TREMBLrel. 05, Created)		
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
GN		POLYSPECIFIC ORGANIC ANION TRANSPORTER.		
GN		OCN1.		
OS		Homo sapiens (Human).		
OC		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OC		NCBI_Taxid=9606;		
RP		[1]		
RP		SEQUENCE FROM N.A.		
RC		TISSUE=LIVER;		
RC		MEDLINE=98086199; PubMed=9426230;		
RA		Tamai I., Yabuuchi H., Nezu J., Sai Y., Oku A., Shimane M., Tsuji A.;		
RT		"Cloning and characterization of a novel human pH-dependent organic		
RT		cation transporter, OCTN1.;"		
RL		FEBS Lett. 419:107-111(1997).		
CC	-1-	SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1-	SIMILARITY: TO SUGAR TRANSPORTER FAMILY.		
DR	EMBL:	AB007448; BAA23356.1; -;		
DR	InterPro:	IPR003662; sub.transporter.		
DR	Pfam:	PF000083; sugar_tr; 1.		
DR	PROSITE:	PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.		
DR	SEQUENCE	551 AA; 62176 MW; F5903421C789F60A CRC64;		

[illegible]

```

QY 181 TGFSLQIFSKNFEMFVFLVGVGMGOISNYVAAAFVLGTEILGKSVRIIFSTLGVCIFFA 240
Db 181 TGFSLQIFSKNFEMFVFLVGVGMGOISNYVAAAFVLGTEILGKSVRIIFSTLGVCIFFA 240
QY 241 FGYMVLPLFAFIRDRWMLLVALTMPGVLCVAVLWVFIPESPRWLSQGRFEAEVIRKA 300
Db 241 VGYMVLPLFAFIRDRWMLLVALTMPGVLCVAVLWVFIPESPRWLSQGRFEAEVIRKA 300
QY 301 AKANGIVVPTIFDPSELQDLSKKQSHNLDLRTNIRMTVIMSIMLWMTISVGYFG 360
Db 301 AKMNTAVPAVIFD--SVEELNPLKQKRAFILDFTFRNIAIMTMSLLMLTSGVYFA 358
QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOYLPRYSMATALFLGSGVLLFMOLV 420
Db 359 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOYLPRYSMATALFLGSGVLLFMOLV 418
QY 421 PPDLYLATVLVVMGKFGVTAAFSMVYVYTAELYPTVVRNMGVGVSTASRLGSLSPYF 480
Db 419 PVDYFELSLGLVLMGKFGVTAAFSMVYVYTAELYPTVVRNMGVGVSTASRLGSLSPYF 478
QY 481 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDITDMLRVKGMKHKRTPSHTR 540
Db 479 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDITDMLRVKGMKHKRTPSHTR 535
QY 541 MKDGOERPTILKSTAF 557
Db 536 DSMTEENPKVL-ITAF 551

```

RESULT 4

```

Q9R141 ID Q9R141 PRELIMINARY; PRT; 553 AA.
AC Q9R141;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ORGANIC CATION TRANSPORTER OCTN1.
GN OCTN1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Wu X., Wang H., Leibach F.H., Ganapathy V.;
RT "Functional characteristics and tissue distribution pattern of OCTN1,
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AF169831; AAD46922.1; -.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 553 AA; 62362 MW; E26C8155768A14AD CRC64;

```

Query Match 74.8%; Score 2157; DB 11; Length 553;
 Best Local Similarity 72.7%; Pred. No. 3.3e-127;
 Matches 405; Conservative 65; Mismatches 83; Indels 4; Gaps 2;

```

QY 1 MRDYDEVTAFLGEGWGFQRLIFLFLSASIIPINGFTGLSSVFLIATPEHRCRVPDANLSS 60
Db 1 MRDYDEVTAFLGEGWGFQRLIFLFLSASIIPINGFTGLSSVFLIATPEHRCRVPDANLSS 60
QY 61 AWRNHTVPLRLRDGREVPHSCRRYRLATIANFSALGLEPGRVDLGLQESCLDGWEYS 120
Db 61 AWRNHTVPLRLRDGREVPHSCRRYRLATIANFSALGLEPGRVDLGLQESCLDGWEYS 120
QY 121 QDYLSTIVTEWNLVCEDDWKPTLTSLFPFVGLGSEFVSQGLSDRFRGKVKLFATMAVQ 180

```

```

Db 121 QDYLSTIVTEWNLVCEDDWKPTLTSLFPFVGLGSEFVSQGLSDRFRGKVKLFATMAVQ 180
QY 181 TGFSLQIFSKNFEMFVFLVGVGMGOISNYVAAAFVLGTEILGKSVRIIFSTLGVCIFFA 240
Db 181 TGFSLQIFSKNFEMFVFLVGVGMGOISNYVAAAFVLGTEILGKSVRIIFSTLGVCIFFA 240
QY 241 FGYMVLPLFAFIRDRWMLLVALTMPGVLCVAVLWVFIPESPRWLSQGRFEAEVIRKA 300
Db 241 VGYMVLPLFAFIRDRWMLLVALTMPGVLCVAVLWVFIPESPRWLSQGRFEAEVIRKA 300
QY 301 AKANGIVVPTIFDPSELQDLSKKQSHNLDLRTNIRMTVIMSIMLWMTISVGYFG 360
Db 301 AKMNTAVPAVIFD--SVEELNPLKQKRAFILDFTFRNIAIMTMSLLMLTSGVYFA 360
QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOYLPRYSMATALFLGSGVLLFMOLV 420
Db 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOYLPRYSMATALFLGSGVLLFMOLV 420
QY 421 PPDLYLATVLVVMGKFGVTAAFSMVYVYTAELYPTVVRNMGVGVSTASRLGSLSPYF 480
Db 421 PVDYFELSLGLVLMGKFGVTAAFSMVYVYTAELYPTVVRNMGVGVSTASRLGSLSPYF 480
QY 481 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDITDMLRVKGMKHKRTPSHTR 540
Db 481 VYLGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDITDMLRVKGMKHKRTPSHTR 537
QY 541 MKDGOERPTILKSTAF 557
Db 538 DSMTEENPKVL-ITAF 553

```

RESULT 5

```

Q9Z306 ID Q9Z306 PRELIMINARY; PRT; 553 AA.
AC Q9Z306;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ORGANIC CATION TRANSPORTER.
GN SLC22A4 OR OCTN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Nezu J.;
RT "Mouse OCTN1 : Polyspecific organic cation transporter.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AB016257; BAA36626.1; -.
DR MGD; MGI:1353479; SLC22a4.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 553 AA; 62290 MW; C4D66BC061398653 CRC64;

```

Query Match 74.5%; Score 2148; DB 11; Length 553;
 Best Local Similarity 72.2%; Pred. No. 1.2e-126;
 Matches 402; Conservative 67; Mismatches 84; Indels 4; Gaps 2;

```

QY 1 MRDYDEVTAFLGEGWGFQRLIFLFLSASIIPINGFTGLSSVFLIATPEHRCRVPDANLSS 60
Db 1 MRDYDEVTAFLGEGWGFQRLIFLFLSASIIPINGFTGLSSVFLIATPEHRCRVPDANLSS 60
QY 61 AWRNHTVPLRLRDGREVPHSCRRYRLATIANFSALGLEPGRVDLGLQESCLDGWEYS 120
Db 61 AWRNHTVPLRLRDGREVPHSCRRYRLATIANFSALGLEPGRVDLGLQESCLDGWEYS 120
QY 121 QDYLSTIVTEWNLVCEDDWKPTLTSLFPFVGLGSEFVSQGLSDRFRGKVKLFATMAVQ 180

```

```
QY 121 ODVYLTIVTEWNLVCEDDWKAPLTLISLFFVGVLLGSGFTSGQLSDRFRGKKNVLFVTMGMQ 180
DB 121 KIDFLSTIVTEWNLVCEDDWKAPLTLISLFFVGVLLGSGFTSGQLSDRFRGKKNVLFVTMGMQ 180
QY 181 TGFSLQIFSKNFMFVFLVGVGQISNYAAVFLGTEILGKSVRIIFSLGVCIEYA 240
DB 181 TGFSLQIFSKNFMFVFLVGVGQISNYAAVFLGTEILGKSVRIIFSLGVCIEYA 240
QY 241 FGYVLPFLFAYFIRDRWMLLVALTMPGVLCVALWMPFIPESPRWLISQGRFEEAEVIRKA 300
DB 241 FGYVLPFLFAYFIRDRWMLLVALTMPGVLCVALWMPFIPESPRWLISQGRFEEAEVIRKA 300
QY 301 ARANGIWPSTFTDSELDQSKKQOSNIIDLRTNWRVWTMSIMLWMTISGVYFG 360
DB 301 ARANGIWPSTFTDSELDQSKKQOSNIIDLRTNWRVWTMSIMLWMTISGVYFG 360
QY 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOVLPFRYSMATATLFGGSLVLFMOIV 420
DB 361 LSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOVLPFRYSMATATLFGGSLVLFMOIV 420
QY 421 PPDLYLATVLMVWGKFGVTAAFSMYVYVTAELYPTVVRNMGVGVSTASRLGSLSPYF 480
DB 421 PPDLYLATVLMVWGKFGVTAAFSMYVYVTAELYPTVVRNMGVGVSTASRLGSLSPYF 480
QY 481 VILGAYDRFLPYLMLGSLTILATLTLPESEFGTLPDIDQMLRVKGMKHKRTPSHR 540
DB 481 VILGAYDRFLPYLMLGSLTILATLTLPESEFGTLPDIDQMLRVKGMKHKRTPSHR 540
QY 541 MLKQOERPTILKSAF 557
DB 541 MLKQOERPTILKSAF 557
QY 538 VSVDRSESPKVL-ITAF 553
DB 538 VSVDRSESPKVL-ITAF 553

RESULT 6
Q9VCA2 PRELIMINARY; PRT: 548 AA.
AC Q9VCA2;
CT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE ORCT PROTEIN.
GN ORCT OR CG6331.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley J.S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
```

```
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relneert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Stradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Welstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; AE003747; AAF56271.1; -.
DR FlyBase; FBgn0019952; Orct.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 548 AA; 61002 MW; 08D7F97599B477AF CRC64;

Query Match 30.7%; Score 885; DB 5; Length 548;
Best Local Similarity 35.1%; Pred. No. 9.9e-48;
Matches 198; Conservative 106; Mismatches 206; Indels 54; Gaps 9;

QY 4 YDEVATFLGEGWGPQRLIFLFLLSASIPNGFTGLSSVFLIATPEHRCRVP-----DAANLS 59
DB 3 YDDVITHLGEFGPYKRIYYLLCLPAIVCAFHKLGLAGVFLAKPDFRCALPYENGSIYELS 62
QY 60 SAWENHTVPLRLDGRVPHSCRRYRLATIANFSLGLEPRDVL-----CQL----- 108
DB 63 PHLWNLSPNER-----CSY-----DVDYTEYLNGSIDRSSN 97
QY 109 EQESCLDGEWFEQSDVYLTSTVTWNLVCEDDWKAPLTLISLFFVGVLLGSGFTSGQLSDRFG 168
DB 98 ETKTC-SSYVYVDRSKYLNSAVTEWNLVCSRLSLSATSDSLFMLGVLLGSLIFGQMSDKLG 156
QY 169 RKNVLFVTMGWQGFSPFLQIFSNFEMFVFLVGVGQISNYAAVFLGTEILGKSVRI 228
DB 157 RKPTFFASLVQLIFGVLAAPAPEYFSTYITRMIVGATTSGVFLVAVIALENVGSYSR- 215
QY 229 IESTLGVCIPIYAFGYVLPFLFAYFIRDRWMLLVALTMPGVLCVALWMPFIPESPRWLISOG 288
DB 216 LFGAVAMQMFSSVGFMLTAGFAYFIRDRWMLLVALTMPGVLCVALWMPFIPESARWILMKG 275
QY 289 RFEEAEVIRKAANGVIVPSTIFD-----PSELQDLSKKKQOSHNLIDLLRTWNR 341
DB 276 RKDEAFVILKEAKENKVEVPNIYEQLVDEVAEKKKQDEMAASQPAATVFDLLRYPNLR 335
QY 342 MYTMSIMLWMTISGVYFGSLDTPNLHGDIFVNCFLSAMVEVPAYVLAWLLOVLPYRY 401
DB 336 RKTLLIFFDFWVNSGVYVGLSWNTNLGGNQLVNFEMISGAVEIPGYTLLLTLLNRGRRS 395
QY 402 SMATALFLGSLVLEFQVLPDLYLATVLMVWGKFGVTAAFSMYVYVTAELYPTVVRNM 461
DB 396 ILCGTMMVAGISLATAIFVPSDMNMLIVACAMIGLAIATSSYGTIYIFSAEQFPTVVRNV 455
QY 462 GVGVSSTASRLGSLSPYFVFLGAYDRFLPYLIMGSLTILATLTLPESEFGTLPDPTI 521
DB 456 GLGASSWAVRGVIGILAPYLKELIWRPLIICGALSITAGLSILLPETLNKMPETI 515
QY 522 DQMLRVKGMKHKRTPSHRMLKDG 545
DB 516 ED-----GENFGKKPAPQETAEEG 534

RESULT 7
O01384
```

ID	NCBI_TaxID=7227;	PRELIMINARY;	PRT;	548 AA.
AC	001384;			
AD	01-JUL-1997 (TREMBLrel. 04, Created)			
AE	01-JUL-1997 (TREMBLrel. 04, Last sequence update)			
AF	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
BT	PUTATIVE ORGANIC CATION TRANSPORTER.			
DE	ORCT OR CG6331.			
GN	Drosophila melanogaster (Fruit fly).			
OS	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OX	Ephydroidea; Drosophilidae; Drosophila.			
OC	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98072431; PubMed=9409773;			
RA	Taylor C.A.M., Stanley K., Shirras A.D.;			
RT	"The Orct gene of Drosophila melanogaster codes for a putative organic			
RT	cation transporter with six or 12 transmembrane domains.";			
RL	Gene 201:69-74(1997).			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.			
DR	EMBL; Y12400; CAA73031.1; -			
DR	FlyBase; Fgn0019952; Orct.			
DR	InterPro; IPR003662; sub.transporter.			
DR	Pfam; PF00083; sugar_tr; 1.			
DR	Transmembrane.			
SK	SEQUENCE 548 AA; 61105 MW; 2665B71C969C5E4A CRC64;			
QY	Query Match	30.6%;	Score 881;	DB 5; Length 548;
DB	Best Local Similarity	34.9%;	Pred No. 1,8e-47;	
DB	Matches 197; Conservative 107; Mismatches 206; Indels 54; Gaps			
QY	4 YDEVTAFLGEMGPPORLIFFLSASIIPINGTGLSGSVFLIATPEHRCRVP---- <td></td> <td></td> <td></td>			
DB	3 YDDVITHLGFEGPYQKRIYLLCLPAIVCAFHKLAVFLAKPDPFCALPYENGSIYELS 62			
QY	60 SAWENHTVPLRLDROGEVPHSCRRYRLATIANFSALGLEPGRDVL-----GQL----- 108			
DB	63 PHLNLSIPNER-----CSYI-----DVDYTEYLNSTIPRSSN 97			
QY	109 EQESCLDGEFSDQVYLSTIVTWNLYCEDDDMKAPLTIISFFVGVLGSLFSGQSDRFG 168			
DB	98 ETKTC-SSYYVDRSKYLSNAVTWNLYSCRSLLSATSDSLFVLGVLLGSLIFGQMSDKLG 156			
QY	169 RKNVLFVTMGQMOTGFELQIFSKNFEMFVVLVFLVCMGQISNYAAVFLUGTEILCKSVRI 228			
DB	157 RKPTFFASVLQILFYGAAVAPEYFYSYIRSMVIGATTSGVFLVAYIALEMVGSSYR- 215			
QY	229 IFSFLGVCIFAFQYVVLFFAFVFIIDRWMLVALTMPGVLCAVLMWFFIPESRWLSQG 288			
DB	216 LFAGVAMQMFSGVFMLTAGFAFIHDWRWLQIATVLPGLFLCYWIIPESARWLLMKG 275			
QY	289 RFEEAEVITIRKAAGKIVVPSTIFD-----PSELQDLSKKQOOSHNLDLLRTWNR 341			
DB	276 RKDEAFVIEKAENKVEVPNEIYEQLVDEVAEKKKQDEMAASQPAATVFOLLRYPNLR 335			
QY	342 MYTMSLMLWMTISGVFGSLDTPNLHGDFIVNCFLSAMVEVPAYVLAWLLOLVLPARY 401			
DB	336 RKTLLIFDFMFVNSGVYIGLSWNTNLGGNQLGNFMFISGVEPIGTYTLFFFTLNRWGRS 395			
QY	402 SMATALFLGGSVLLFMQLVDPDLVLYATVLVMVGKGVTAAFSMVYVYAEVLYPWWNM 461			
DB	396 ILCGTMVAGISLTLTFVPSDMNWLIVACAMIGKLAITSSTVGTIYFSAEQFPVVRNV 455			
QY	462 GVGYSSTASRLGSLSPYFVYLGAIDRFPLIYILMGSLTILTLTLFLPESFGTLPDTI 521			
DB	456 GLGASSVAVRGVIGLAPYLKLLGEIWRPLIICGALSITAGLLSRLLPETLNKMPETI 515			
QY	522 DQMLRVGKMKHRKTPSTRMLKG 545			
DB	516 ED-----GENFGKKPAQPAETAE 534			

```

RESULT      8
Q9VCA3
ID          PRELIMINARY;      PRT;      567 AA.
AC          Q9VCA3;
DT          01-MAY-2000 (TEMBLrel. 13, Created)
DT          01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT          01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE          CGI3610 PROTEIN.
GN          CGI3610.
GS          Drosophila melanogaster (Fruit fly).
OC          Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC          Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC          Ephydroidea; Drosophilidae; Drosophila.
OR          NCBI_TaxID=7227;
RX          [1]
SEQUENCE FROM N.A.
PC          STRAIN=BERKELEY;
EX          MEDLINE=20196006; PubMed=10731132;
RA          Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA          Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA          George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA          Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA          Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA          Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA          Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA          Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA          Beeson K.Y., Bencos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA          Borius K.C., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA          Burkova D., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA          Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA          de Pablos B., Delcher A., Deng Z., Dugan-Rochet S., Dietz S.M.,
RA          Dodson K., Doup L.E., Downes M., Dugan-Rochet S., Dunkov B.C., Dunn P.,
RA          Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA          Foster C.G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA          Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA          Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA          Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA          Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA          Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA          Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA          Liu X., Mattet B., McIntosh A.C., McLeod M.P., McPherson D.,
RA          Merkulov G., Milshina N.V., Mobarry L., Murphy L., Muzny D.M., Nelson D.L.,
RA          Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA          Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA          Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA          Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA          Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA          Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA          Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA          Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA          Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA          Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA          Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA          Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT          "The genome sequence of Drosophila melanogaster.";
RL          Science 287:2185-2195(2000).
CC          -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC          -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR          EMBL; AE003747; AAF56270.1; -
DR          FlyBase; FBgn0039176; CGI3610.
DR          InterPro; IPR003662; sub_transporter.
DR          Pfam; PF00083; sugar_tr; 1.
KW          Transmembrane.
SQ          SEQUENCE      567 AA;      63166 MW;      0D9B979469395E60 CRC64;

```

Query Match	28.3%	Score 816;	DB 5;	Length 567;	...
Best Local Similarity	32.2%	Pred. No. 2.1e-43;			
Matches 183;	Conservative 121;	Mismatches 212;	Indels 52;	Gaps 9	

QY 4 YDEYTAFLGEWGPFORLIFFLLSASIIPNGFTGLSSVFLIATPEHRCRVPDAANLSSAWR 63

Db 3 YDEAIHLGDFGRYQKIIVFLICLTPSIPVAFHKLAGVFLAKPDRFCALPFENGSSYDLP 62
 QY 64 NHTVPLRLRDGREGVPHSCRRYRLATIANFSALGLEPGRDVL-----GQL-----ESES 112
 Db 63 THLWNLSPYENER-----CSY-----DVTTEYLNGSIPRNSNETKT 101
 QY 113 CLDGEWFSQDVYLSITVEMNLVCEDDWKAPLTISLFFVGVLLGSFISQSLSDREGRKNV 172
 Db 102 C-SSYVYDRSKYLSAVTEWNLVCGRDPMAATSDSLFMLGVLLGSFISQSLSDREGRKNV 160
 QY 173 LFTVMGQWQFGLQIFSKNFEMFVFLVLCMGQISNVAAVFLGTEILGKSVRIIFST 232
 Db 161 LFASLVIQVLFGLAGVAPEYFTYTFARLMVGATTSGLVLAIVVAMVMGPKR-LVAG 219
 QY 233 LGVCIFYAFGYVVLPLFAVFIQWRMLVALTMPGVLCVLAWEIPESPRWLISOGREE 292
 Db 220 IFVMEFFSGVGLMATAVFAVFDHWRQLALPLGLIIFWYIIPESARWLLKGRDC 279
 QY 293 AEVIIRKAAKANGIVP-----STIFDPS-----LQDLSKKQKSHNILDLLRTWN 339
 Db 280 ATANMOKAARFNKVEISDEALSELLDEGENSEKAKQLEQDELDEGPPSWMDLFCYPN 339
 QY 340 IRWVITMSIMLWMTISVGVGLSLDTPNLHGDFVNCFLSAMVEPAYVLAWLILQYLPR 399
 Db 340 LRKTLIIFLDLWVTSVGVYGLSWNTSLGGNVLLNFVISGAVEIPAYIFILLTLNRWR 399
 QY 400 RYSMTALFLGGSVLEFMQVLPDLYLATVLMVVGKFGVTAAFSMVYVYTAELYPTVVR 459
 Db 400 RSLGCLVMAGLSLATVLIIPQRMHTLIVACMLGKLAITASYGTVIFSAEQFPTVVR 459
 QY 460 NMGVGSSTASRLGSLSPYFVILGAYDRFLPYILMGSLTILTAITLFLPESFGTLPD 519
 Db 460 NYALGAASVARISSQMAPFLNATLWPLPLGCLSLTILVAGLSLLPLPETHNKPML 519
 QY 520 TIDQMLRVKGMKHKRTPSHTRMLKDGQE 547
 Db 520 TI-----ADGERFGKTRADVLETGQE 542

RESULT 9
 Q90539
 ID Q90539 PRELIMINARY; PRT; 568 AA.
 AC Q90539
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ORGANIC CATION TRANSPORTER 1.
 GN OCT-1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99227113; PubMed=10209228;
 RA Wu X., Fei Y.J., Huang W., Chancy C., Leibach F.H., Ganapathy V.;
 RT "Identity of the F52F12.1 gene product in Caenorhabditis elegans as an
 RL organic cation transporter.";
 RL Biochim. Biophys. Acta 1418:239-244(1999).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC EMBL; AF110415; AAF21932.1;
 DR InterPro; IPR003662; sub_transporter.
 DR Pfam; PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 568 AA; 63505 MW; COA3E73851F44056 CRC64;

Query Match 27.08; Score 777; DB 5; Length 568;
 Best Local Similarity 31.4%; Pred. No. 5.8e-41;
 Matches 165; Conservative 118; Mismatches 225; Indels 18; Gaps 6;
 QY 3 DYDEVTAFLGEGWGFQRLIFLFLS-ASIIIPNGFTGLSVFLIATPEHRCRVPDAANLSSA 61

Db 8 DDFVLEQVNGYQIYVFFFIICLTPSLPSAFSAFNIPFVGNPPHTCHIPGKEYLRP 67
 QY 62 WRNVTPLRLRDGREGVPHSCRRYRLATIANFSALGLEPGRDVLGQLEQSLDGEWFSQ 121
 Db 68 LTNDQTL-----SCKQNETQINVFRAFTSAP-VDTYSDRISLVPCCQNGWDYDN 116
 QY 122 DVYLSITVTEWNLVCEDDWKAPLTISLFFVGVLLGSFISQSLSDREGRKNVLFVWGMQT 181
 Db 117 STYLDLSVTFNLCVDOQAQIEISTTSFYVGSFIGNCLFGYVADKFGRRRSFFVILT 176
 QY 182 GFSLQIFSKNFEMFVFLVLCMGQISNVAAVFLGTEILGKSVRIIFSTGLGCFYAF 241
 Db 177 VCGFASFAKDIIESFIILREFTGLAPALQIPFIICMEFGNSGR-IFSGLMSTLFFGA 235
 QY 242 GYVVLPLFAVFIQWRMLVALTMPGVLCVLAWEIPESPRWLISOGREEAEVIRKAA 301
 Db 236 AMALLGVVAMFIRWRQLTFFCNAPFAFYIYFFELPESPRWSVGVKADAKKQLKIA 295
 QY 302 KANG--IVVPTSTIFDPSLQDLSKKQ--QSHNILDLLRTWIRWVITMSIMLWMTISV 356
 Db 296 KMGKSNVDDELVDKSNHONAAEETKRSNVTDLFKTPNLRKTLIVTIWYNNAI 355
 QY 357 GYGLSLDTPNLHGDFVNCFLSAMVEPAYVLAWLILQYLPRYSMTALFLGGSVLLF 416
 Db 356 IYNGLTENSNLPLVDDYWSFIINGAVELPGYFVVMPLQACAGRRWTLAATMIVCGIGCVS 415
 QY 417 MOLVPPDLYLATVLMVVGKFGVTAAFSMVYVYTAELYPTVVRNMGVSVSTASRLGSL 476
 Db 416 AMFMDGYPMLVASASFIKFGVSGFAVIYIFAGELYPTVVRNMGVSVSTASRLGSL 475
 QY 477 SPYFVILGAYDRFLPYILMGSLTILTAITLFLPESFGTLPDID 522
 Db 476 APLVNLGKIVKILPLLMGLMALSAGILFFLPLGAPLPMTIE 521

RESULT 10
 O02270
 ID O02270 PRELIMINARY; PRT; 576 AA.
 AC O02270
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE F52F12.1 PROTEIN.
 GN F52F12.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matthews L.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
 RA Craxton M., Dear S., Du Z., Durbin R., Pavello A., Fulton L.,
 RA Gardner A., Green P., Hawkins J., Hillier L., Jier M., Johnston L.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P., III of C.
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Nature 368:32-38(1994).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
 DR EMBL; Z83228; CAB05732.1;
 DR InterPro; IPR003662; sub_transporter.

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottiier P.
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahle C., Daverton L.B., Davies P.,
de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I.J., Dietz S.M.,
Dodson K., Doug L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleisschmann W.,
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jajali M.K., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Lin X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosherfi A.,
Mount S.M., Murphy B., Murphy L., Murzy D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nussken R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.K., Sidon-Klimas I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Teclor C., Turner R., Ventner E., Wang A.H., Wang X.,
Wang Z.-Y., Wasserman D.A., Weinstein G.W., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.

"The genome sequence of *Drosophila melanogaster*.":
Science 287:2185-2195(2000).
RL
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
EMBL; AE003819; AAC36407.1; -.
DR FlyBase; FBgn0033809; CG4630.
DR InterPro; IPRO03662; sub_transporter.
DR pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 561 AA; 62041 MW; 573643481595788F CRC64;

Db 420 GYSLAWLFLRRGRRVALSGSLILCSITCVASGFTVLGANLWLVTLFLVGKLGITSSRAV 479

QY 446 VYVYTAELPTVVRNMGVSVSTASRLGSLSPYFVYLGAYDRFLPYILMGSLTILTAIL 505

Db 480 IYFTAEEMPTVIRSGGVGMSTFARFGAMLAPFVPLASYYDPLPLLLFGTSLVAGIL 539

QY 506 TLPSPESFGTLPDPDT 520

Db 540 SLLLPETFNKLPDPDT 554

RESULT 12

Q9NQD4 PRELIMINARY; PRT; 554 AA.

AC Q9NQD4

DT 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)

DE PARTIAL OCT1 ORGANIC CATION TRANSPORTER, EXON 1 AND JOINED CDS.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Hayer M., Bonisch H., Bruss M.;

RT "Molecular cloning, functional characterization and genomic

RT organization of four alternatively spliced isoforms of the human

RT organic cation transporter 1 (hOCT1/SLC22A1).";

RL Ann. Hum. Genet. 63:473-482(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA Hayer M., Bonisch H., Bruss M.;

RT "Molecular cloning, functional characterization and genomic

RT organization of four alternatively spliced isoforms of the human

RT organic cation transporter 1 (hOCT1/SLC22A1) (corrigendum).";

RL Ann. Hum. Genet. 64:267-267(2000).

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

DR EMBL; AJ243995; CAB95971.1; JOINED.

DR EMBL; AJ243996; CAB95971.1; JOINED.

DR EMBL; AJ276051; CAB95971.1; JOINED.

DR EMBL; AJ276052; CAB95971.1; JOINED.

DR EMBL; AJ276053; CAB95971.1; JOINED.

DR EMBL; AJ245460; CAB95971.1; JOINED.

DR EMBL; AJ243998; CAB95971.1; JOINED.

DR EMBL; AJ243999; CAB95971.1; JOINED.

DR EMBL; AJ244000; CAB95971.1; JOINED.

DR InterPro; IPR003662; sub-transporter.

DR Pfam; PF00083; sugar_tr; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.

SW Transmembrane.

KW SEQUENCE 554 AA; 61169 MW; 9CA1A6C0FC2B2A96 CRC64;

Query Match 25.7%; Score 740; DB 4; Length 554;

Best Local Similarity 35.9%; Pred. No. 1.2e-38;

Matches 197; Conservative 86; Mismatches 215; Indels 50; Gaps 12;

QY 1 MRDYDEVTAFLGEMGPFOR---LIFLLASATIIPNGFTGLSSVELIATPEHRCRVPOAAN 57

Db 1 MPTVDDILEQVGESGWFOQAFLICLLUSAAP---ICVGIVELGFTPDHHCOSPGVAE 57

QY 58 LSS---AWR-----NITVPLRLRDGVEPHSCRRYR-----LATIA-NFSALG 96

Db 58 LSQRGWSPAEELNTPVGLPAGEAFIQCRRYEDVQNSQSLSCVDPPLASLATNRSHLP 117

QY 97 LEPGRDVLGQLEQSGLDGNEFSQDVLSTIVTEWNLVCEDDWKAPLITSLFFVGVLLG 156

Db 118 LGP-----CDQGVY---DTPGSSIVTEFNLCVADSQKDLDFQSLNAGELFG 162

QY 157 SPISQSLDRGRKRVLFVMTGMOTGFSFLQIFSKNEMFVFLVCMGQISNVAAVF 216

Db 163 SLGVGYFADRFGRKLCILGTVLVNNAVSGVLMASPNTVMSMLLFRLLQGLVSKGNWAGYT 222

QY 217 LGTEILKSVRIIPSTLGVCFYAF--GYMVLPLFAYFIRDWRMLLVALTMPGVLCVALW 274

Db 223 LITEFVGSGSR--RTVAIMYQMAFTVGLVALTGLAVALPHRWQLAVSLPTFLFLYY 279

QY 275 WFIPESRWLLISQORFEAEVIRKAAKANGIVVPSITFDPSLQDLSSKKQSHNLDL 334

Db 280 WCVPESPRWLLSKQRNTEAIKIMDHIAQKNGKLPADLKLMSLEEDVTER--LSPSFADL 337

QY 335 LRTWNIRMTIMSLMWTISVGVEGLSLDTPNLHGDFVNCFLSAMVEVPAYVLAWL 394

Db 338 FRTLRLKRTFILMFTSVLQGLLHMGTSGNLYLDFLYSALVEIPGAFIALITI 397

QY 395 QYLPRRYSMATLFLGGSVLLFMQVPPDYLYLATVLVMVGKFGVTAAFSMVYVYTAEL 454

Db 398 DRVGRIYPMAMSNLLAGAACLVIMIFISPDHLNLNIIIMCVGRMGITTAIQMICLVNAELY 457

QY 455 PTVVRNMGVSVSTASRLGSLSPYFYV--LGAYDRFLPYILMGSLTILTAILFLPESF 513

Db 458 PTFVRNLGVWVCSSLCDIGGIITPFIVFRLEVRVQWQLPLILFAVLGLLAAGVTLILPETK 517

QY 514 GTPLPDTI 521

Db 518 GVALPETM 525

RESULT 13

Q75751 PRELIMINARY; PRT; 556 AA.

AC Q75751; Q9UP02;

DT 01-NOV-1998 (TReMBLrel. 08, Created)

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)

DE ORGANIC CATION TRANSPORTER 3 (EXTRANEURONAL MONOAMINE TRANSPORTER)

DE (EMT) (SOLUTE CARRIER FAMILY 22, MEMBER 3).

GN SLC22A3 OR EMT.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

RC TISSUE=KIDNEY;

RX MEDLINE=99212254; PubMed=10196521;

RA Gruendemann D., Schechinger B., Rappold G.A., Schoemig E.;

RT "Molecular identification of the corticosterone-sensitive

RT extraneuronal catecholamine transporter.";

RL Nat. Neurosci. 1:349-351(1998).

RN [2]

RP SEQUENCE OF 513-556 FROM N.A., AND TISSUE SPECIFICITY.

RX MEDLINE=99134299; PubMed=9933568;

RA Verhaagh S., Schweifer N., Barlow D.P., Zwart R.;

RT "Cloning of the mouse and human solute carrier 22a3 (SLC22a3/SLC22A3)

RT identifies a conserved cluster of three organic cation transporters on

RT mouse chromosome 17 and human 6q26-q27.";

RL Genomics 55:209-218(1999).

RN [3]

RP FUNCTION, AND TISSUE SPECIFICITY.

RC TISSUE=KIDNEY;

RX MEDLINE=20425388; PubMed=10966924;

RA Wu X., Huang W., Ganapathy M.E., Wang H., Kekuda R., Conway S.J.,

RA Leibach F.H., Ganapathy V.;

RT "Structure, function, and regional distribution of the organic cation

RT transporter OCT3 in the kidney.";

RL Am. J. Physiol. 279:F449-F458(2000).

CC -!- FUNCTION: MEDIATES POTENTIAL-DEPENDENT TRANSPORT OF A VARIETY OF

CC ORGANIC CATIONS.

CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

CC -!- TISSUE SPECIFICITY: EXPRESSED IN PLACENTA, SKELETAL MUSCLE,

CC PROSTATE, AORTA, LIVER, FETAL LUNG, SALIVARY GLAND, ADRENAL GLAND,

CC KIDNEY AND BRAIN CORTEX. NO EXPRESSION DETECTED IN SPLEEN.

CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC

```
CC CATION SUBFAMILY.
DR EMBL; AJ001417; CA004751.1; -.
DR EMBL; AF078749; AAD20977.1; -.
DR MIM; 604842; -.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
KW Transport; Ion transport; Transmembrane; Glycoprotein.
FT TRANSMEM 21 41 POTENTIAL.
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 241 261 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.
FT TRANSMEM 468 488 POTENTIAL.
FT TRANSMEM 498 518 POTENTIAL.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 556 AA; 61279 MW; C3CA2D77DD21C658 CRC64;

Query Match 25.6%; Score 738.5; DB 4; Length 556;
Best Local Similarity 33.0%; Pred. No. 1.4e-38;
Matches 189; Conservative 117; Mismatches 214; Indels 53; Gaps 18;

QY 1 MRDYDEVTAFLGEGPQRLIFELLSSAIPNGFTGLSSVFLIATPEHR-CRYPDAANLS 59
DB 1 MPSDEALQVGEFGFRQRRVFLLLCTGTFAFLVGVVFLGTQPDHYWCGRPSAALA 60
QY 60 S--AWR-----NHTVPLRLRDGREVPH---SCRRYRLATIANFSA-----LGLE 98
DB 61 ERCWSPPEEWNTAP--ASRGPEPERRRCQRY--LLEANDSASATSAISCADPLAF 117
QY 99 PGRDVLQGLEQESCLDGEFSQDVYLSITVTENLVCEDDWKAPLTISLFFVGVLLG 158
DB 118 PNRSAPL-----VPCRGGRYAO--AHSTIVSEFDLCVNAWMLDQTQAILNLGFLTGF 170
QY 159 ISGQLDRGRKKNVLFVTMGQMTGFSLQI-FSKNFEMFVFLVGMGQISNYAAFLV 217
DB 171 TLGYAADRYGR-IVYLLSLGIVGTGVVVAFAFAPFPVFIFRFLQGVFGKGTWTCYVI 229
QY 218 GTEILGKSVRIIFSTLGVCI--FVAFGYMVLPLFAFYFIRDRMMLLVALTMPGVLCA 275
DB 230 VTELVGSKQRI---VGIVIQMFTEIILPGIAIFIPNWQIQALITLPSFLFLCYW 286
QY 276 FIPESPRWLISQGRFEAEVIRKAANGIVVPSTIFDPSQLDSSKKQQSHNILDLL 335
DB 287 VWPESPRWLITRKKGKALQILRIAKCNKYLSSNY---SEI-TVTDEVSNPFLDLV 342
QY 336 RTWNRIMVTIMSLMWTISVGFGLSDPNLHGDIFVNCFLSAMVEVPAYVLAWLILQ 395
DB 343 RTPQRRKCTILMFANFTSAVYQGLVMRLGIIGNLYIDFISGVVLPGLALLITIE 402
QY 396 YLPYRYSMATALFGSVLLFMQLVPPDLYLATVLMVKCFGTAAFSVYVYTAELXP 455
DB 403 RLGRRLPFAASNIVAGVACLVTAFLPGIAWLRTVTATLGRIGITWAFELVLYVNSLXP 462
QY 456 TVVRNMGVSVSTASRLSGLSILSPFYV--LGAYDRFLPYILMGSLTITAILTLFLPESFG 514
DB 463 TTLNFGVSLCSGLCDFGGIIAPFLFLRLAAVWLELPILIFGILASICGGLVMLLPETKG 522
QY 515 TPLPDTIDQMLRV-----KGMKHKTP-----SH 538
DB 523 IALPETVDDVEKLGSPHSCRCGRNKKTPVSRSH 555

RESULT 14
O15395 PRELIMINARY; PRT; 554 AA.
AC O15395;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
```

```
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ORGANIC CATION TRANSPORTER 1.
GN HOC1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97330822; PubMed=9187257;
RA Zhang L., Dresser M.J., Gray A.T., Yost S.C., Terashita S.,
RA Giacomini K.M.;
RT "Cloning and functional expression of a human liver organic cation
RL transporter";
RL Mol. Pharmacol. 51:913-921(1997).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL; U77086; AAB67703.1; -.
DR InterPro; IPR003662; sub.transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Transmembrane.
SQ SEQUENCE 554 AA; 61153 MW; 55206B897DE32202 CRC64;

Query Match 25.6%; Score 738; DB 4; Length 554;
Best Local Similarity 35.9%; Pred. No. 1.5e-38;
Matches 197; Conservative 86; Mismatches 215; Indels 50; Gaps 12;

QY 1 MRDYDEVTAFLGEGPQRLIFELLSSAIPNGFTGLSSVFLIATPEHR-CRYPDAAN 57
DB 1 MPTVDILQEVGESWFOKQAFILICLLSSAFAF---ICVGVFLGFTPDHHCQSPGVAE 57
QY 58 LSS--AWR-----NHTVPLRLRDGREVPHSCRRYR-----LATIA-NFSALG 96
DB 58 LSQRGWSPAEELNVTFCIGPAGEAFLQCRRRYEVDNQSALSCVDPPLASLATNRSLP 117
QY 97 LEPGRDVLQGLEQESCLDGEFSQDVYLSITVTENLVCEDDWKAPLTISLFFVGVLLG 156
DB 118 LGP-----CQDGVWY--DTPGSSIVTEFNLVCADSWKLDLFSCLNAGFLFG 162
QY 157 SFISGQLDRGRKKNVLFVTMGQMTGFSLQI-FSKNFEMFVFLVGMGQISNYAAAFV 216
DB 163 SLGVGYFADRFGRKCLLGLTVLVNAVSGVLMAFSNYMSMLLFRLLQLGLVSKGNMAGYT 222
QY 217 LGTEILGKSVRIIFSTLGVCIYAF--GYMVLPLFAFYFIRDRMMLLVALTMPGVLCA 274
DB 223 LITEFVSGSR---RTVAIMYQMAFTVGLVLTGLAYALPHWRWLQAVSLPTFLFLYY 279
QY 275 WPIESPRWLISQGRFEAEVIRKAANGIVVPSTIFDPSQLDSSKKQQSHNILDLL 334
DB 280 WCVPESPRWLISQGRFEAEVIRKAANGIVVPSTIFDPSQLDSSKKQQSHNILDLL 337
QY 335 LRTNIRMTVIMSIMLWMTISVGFGLSDPTNLDGDFVNCFLSAMVEVPAYVLAWLIL 394
DB 338 FTPLRLKRTFLMTLWTFDTSVLYOGLIHMGATSGNLYLDLFLYSALVEIFCAITALITI 397
QY 395 QYLPYRYSMATALFGSVLLFMQLVPPDLYLATVLMVKCFGTAAFSVYVYTAELXP 454
DB 398 DRVGRIYPMAMSNLLAGAACLVWIFISPDHLWLNIIIMCVGRMGITIAIQMCLVNAELY 457
QY 455 PTWNRNMGVSVSTASRLSGLSILSPFYV--LGAYDRFLPYILMGSLTITAILTLPLPESF 513
DB 458 PTFVRLNGLVMVGSLSICDGGIITPFIIVFRRLREVWQALPILFAVLGLLAAGVTLLLPETK 517
QY 514 GTPLPDTI 521
DB 518 GVALPETM 525

RESULT 15
O15245
```

Search completed: January 25, 2002, 10:14:56
Job time: 337 sec

```
ID O15245 PRELIMINARY; PRT; 554 AA.
AC O15245;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ORGANIC CATION TRANSPORTER.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=97405886; PubMed=9260930;
RA Gorboulev V.G., Ulzheimer J.C., Akhondova A.;
RT "Cloning and characterization of two human polypeptide organic cation
RT transporters.";
RL DNA Cell Biol. 16:871-881(1997).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO SUGAR TRANSPORTER FAMILY.
DR EMBL: X98332; CAA66977.1;
DR InterPro: IPR003662; sub.transporter.
DR Pfam: PF00083; sugak_tr; 1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_2.
KW Ion transport; Transmembrane.
SQ SEQUENCE 554 AA; 61187 MW; B16E200852CC4000 CRC64;

Query Match 25.58; Score 734; DB 4; Length 554;
Best Local Similarity 35.88; Pred. No. 2.7e-38;
Matches 196; Conservative 86; Mismatches 216; Indels 50; Gaps 12;

QY 1 MRDYDEVTAFLGEMGFQK---LTFELLSASITPNNGFTGLSSVFLIATPEHRCRVPPDAAN 57
DB 1 MPTVDILLEQVGESWFOKQAFLLICLLLSAFAF---ICVGIYFLGTPDHCQSPGVAE 57

QY 58 LSS---AWR-----NHTVPLRLDRGVPHSCRRYR-----LATIA-NFSALG 96
DB 58 LSQRCGWSFAEELNTYVPLGLPAGEAFLGQCRRYEVDWNSQALSVDPLASLATNRSHLP 117

QY 97 LEPGRDVLGQLEQSCLDGWEFSQDVYLSITVTENLVCEDDKAKPLTISLFFVGVILG 156
DB 118 LCP-----CQDGVY--DTPGSSIVTEFNLCADSWKLDLFQSLNAGFFFG 162

QY 157 SPISQLSDRFGKRVLEVTMGMTGFSFLQIFSKNFEMFVFLVYGMQISNYVAAY 216
DB 163 SLGVGYFADRFGRKICLLGLTVLNAVSGVLMAFSPNYMSMLLFRLLQGLVSKGNMAGYT 222

QY 217 LGTEILGKSVRIIFSTLGCIFYAF--GYMVLPLFAFYFIRDRWMLLYALTMPGVLCVALW 274
DB 223 LITEFVGSGR---RTVAIMYOMAFVTLGLVALTGLAYALPHRWLQAVLSLPTFLFLYY 279

QY 275 WPIPSRWLISQGFEEAEVIRKAKANGIVVPSTIFDPSLQDLSSKKQOQSHNIDL 334
DB 280 WCVPEPSRWLISQKNTAIAKIMHIAQKNGKLPADLKMLSEEDVTEK--LSPSFADL 337

QY 335 LRTWNIRMTIMSIMLMWTISVGYFGLSDPPLNHLGDIYFNCFLSAMVEPAYVYLAWL 394
DB 338 FRTPLRKRTFILMYLWFTDSVLYOGLLHMGATSGNLYLDLYSALVPEIPGAFIALITI 397

QY 395 QYLPRYSNATALFGLGSLVLFMLQVPPDLYLATVLMVGKFGVTAAFSVMYVYTAELY 454
DB 398 DRVGRIYPMAMNLLAGAACLVMIFFISPDHLWLNIIIMCVGRMGITTAIQMICLVNAELY 457

QY 455 PTVVRNMGVGSSTASRLGSILSPFYVY-LGAYDRELPIYILMGSLTILTALTILFPESE 513
DB 458 PTFVRLNMGVWCSSICDGGIITFEIVRELWQALPLILFAVLGLAAGVTLLLPETK 517

QY 514 GTPLPDTI 521
DB 518 GVALPETM 525
```